MODIS Land Cover Product
Algorithm Theoretical Basis Document (ATBD)
Version 4.1

MODIS Land Cover and Land-Cover Change Principal Investigators:

Alan Strahler
Center for Remote Sensing
Department of Geography
Boston University
Boston, MA

John Townshend
Department of Geography
University of Maryland, College Park
College Park, MD

Document Prepared By

Alan Strahler,[1] John Townshend,[2] Douglas Muchoney,[1]
Jordan Borak,[1] Mark Friedl,[1] Suchi Gopal,[1] Andrew Hyman,[1]
Aaron Moody,[3] and Eric Lambin[4]

[1]Boston University
[2]University of Maryland, College Park
[3]University of North Carolina, formerly of Boston University
[4]Université Catholique du Louvain

1 December 1996

Abstract

This document details the structural definition, development process, and functional flow of the MODIS Land-Cover Product. The Land Cover and Land-Cover Change Parameters were proposed by the MODIS Land Team, with Team Member Alan Strahler leading the effort. The Land Cover Parameter is a 1-km product provided on a quarterly basis beginning one year following the launch of the MODIS instrument aboard the EOS-A platform in June, 1998. The Land-Cover Change Parameter is a post-launch, near-term parameter also planned for quarterly delivery.

Both parameters rely on a 1-km gridded database composited from MODIS Level 2 and 3 products. Inputs include (1) surface reflectances at 1000-m spatial resolutions in the MODIS Land Bands (1-7) (an interim product provided by MODIS BRDF/Albedo Product MOD43); (2) spatial texture derived from Band 2 (near-infrared) at 250-m resolution; (3) terrain elevation information (MOD03); (4) vegetation index (MOD34); (5) MODIS Level 3 Snow/Sea Ice cover (MOD33); (6) land surface temperature at 1 km (MOD11; 2484); and (7) directional reflectance information derived from the MODIS BRDF/Albedo Product (MOD43; 3669). These data are composited over a one-month time period to produce a globally-consistent, multitemporal database on a 1-km grid as input to classification and change characterization algorithms.

The Land Cover Parameter recognizes 17 categories of land cover following the scheme proposed by the IGBP for the global 1-km vegetation database, which is now being prepared from global AVHRR LAC NDVI composites. This set of cover types includes eleven categories of natural vegetation covers broken down by life form, three classes of developed and mosaic lands, and three classes of non-vegetated lands.

Land cover classes are assigned by processing the 32-day database using a statistical classification algorithm, *e.g.*, an artificial neural network classifier, to assign land cover classes based on training data. To reduce computational overhead and increase flexibility, classification proceeds by continents.

The Land-Cover Change 1-km Parameter is designed to quantify subtle and progressive land-surface transformations, *i.e.*, land cover modifications, as well as obvious and instantaneous changes, such as land cover conversions. As such, it is not a conventional change product that simply compares land cover databases at two different times and identifies changes in categorical land cover. The algorithm for the Land-Cover Change Parameter combines analyses of change in multispectral-multitemporal data vectors with models of vegetation change mechanisms to recognize both the type of change as well as its intensity. The Land-Cover Change 250-m Parameter has the objective to locate areas of change using the finest resolution MODIS data.

The algorithm development and validation efforts for the Land Cover Product are based on a network of test sites chosen to represent major global biomes and cover types. Five types of sites are recognized, ranging from a few data-rich sites to a larger number

of sites for which only minimal data are available. Prelaunch efforts will focus on sites for which temporal sequences of TM and AVHRR data are available, coupled with fine resolution cover class data. These are most suitable for algorithm development. Post-launch, all types of sites will be used to train the classifier and validate its output. Landsat-7 and ASTER images will be particularly useful for updating information about test sites and identifying local change processes. The validation procedure will characterize the accuracy of the product as well as provide information that can be used in spatial aggregation to provide land cover and land-cover change data at coarser resolutions.

Present schedules call for the first Land Cover Parameter to be released about 15 months after launch based on at least one year of 32-day composites. During this interim period, a provisional at-launch product will be provided. The first Land-Cover Change Parameter will be released about 27 months after launch, based on a two-year period of observation.

New Team Member John Townshend has proposed two new postlaunch products. One of these focuses on an “alarm” algorithm applied to 250-m data that recognizes land cover change, especially deforestation, using changes in spatial texture. The other provides continuous fields of land cover properties, such as proportion in woody cover, using linear mixture modeling.

Table of Contents

Tables

MODIS Land Cover Product
Algorithm Theoretical Basis Document
Version 4.1

1. Introduction

1.1 Identification

MODIS Product No. 12 (MOD12)			
Parameter Number	*Parameter Name*	*Spatial Resolution*	*Temporal Resolution*
2669	Land Cover Type, 1-km	1 km	4/yr
TBD	Land Cover Type, CMG	1/4°	4/yr
2671	Land-Cover Change, 1-km	1 km	4/yr
TBD	Land-Cover Change, 250-m	250 m	TBD

1.2 Overview

This document details the structural definition, development process, and functional flow of the MODIS Land-Cover Product. It represents a revision of the Version 3.0 (Strahler *et al*., 1995) and Version 4.1 (provisional) Land Cover ATBDs. The Land Cover Parameter is a 1-km product provided on a quarterly basis beginning one year following the launch of the MODIS instrument aboard the EOS-AM1 platform in June, 1998 (Running *et al*., 1994b). It will also be prepared on a 1/4^{o}grid for use by global modelers. The Land-Cover Change 1-km Parameter is a post-launch, near-term parameter. The Land-Cover Change 250-m Parameter, proposed by new MODIS Land Team member John Townshend, will be produced at, or very soon after, launch.

The Land Cover and Land-Cover Change 1-km Parameters rely on a 1-km gridded database assembled from MODIS Level 3 products produced on 8- or 16-day cycles. Inputs include:

(1) Surface reflectances derived from the MODIS BRDF/Albedo product (MOD43) in the MODIS Land Bands (1-7), atmospherically corrected and adjusted to nadir view at the prevailing sun angle of the period;
(2) Spatial texture derived from Band 2 (near-infrared, 250-meter) at 1000-m resolution;
(3) Terrain elevation information (MOD03);
(4) Vegetation index (MOD34);
(5) Snow/Sea Ice cover (MOD33);
(6) Land surface temperature at 1 km (MOD11; 2484);
(7) Directional reflectance information from MOD43; and a

(8) Land/water mask that restricts classification to land regions derived from the MODIS BRDF/Albedo Product (MOD43; 3669).

The temporal resolution of this database is 32 days, corresponding to two complete MODIS orbit cycles.

The Land Cover Parameter recognizes 17 categories of land cover following the scheme proposed by the IGBP for the global 1-km vegetation database, which is now being prepared from global AVHRR LAC NDVI (Advanced Very High Resolution Radiometer, Local Area Coverage, Normalized Difference Vegetation Index) composites. This set of cover types includes eleven categories of natural vegetation covers broken down by life form, three classes of developed and mosaic lands, and three classes of non-vegetated lands.

Land cover classes are assigned by processing the 32-day database using a statistical classification algorithm, *e.g.*, an artificial neural network classifier, to assign land cover classes based on training data. To reduce computational overhead and increase flexibility, classification proceeds by continents.

The Land-Cover Change 1-km Parameter is designed to quantify subtle and progressive land-surface transformations, *i.e.*, land cover modifications, as well as obvious and instantaneous changes, such as land cover conversions. As such, it is not a conventional change product that simply compares land cover databases at two different times and identifies changes in categorical land cover. The algorithm for the Land-Cover Change Parameter combines analyses of change in multispectral-multitemporal data vectors with models of vegetation change mechanisms to recognize both the type of change as well as its intensity. The Land-Cover Change 250-m Parameter has the objective to locate areas of change using the finest resolution MODIS data.

The algorithm development and validation efforts for the Land Cover Product are based on a network of test sites chosen to represent major global biomes and cover types. Five types of sites are recognized, ranging from a few data-rich sites to a larger number of sites for which only minimal remote sensing and collateral data are available. Prelaunch efforts will focus on sites for which temporal sequences of TM and AVHRR data are available, coupled with fine resolution cover class data. These are most suitable for algorithm development. Postlaunch, all types of sites will be used to train the classifier and validate its output. Landsat-7 and ASTER images will be particularly useful for updating information about test sites and identifying local change processes. The validation procedure will characterize the accuracy of the product as well as provide information that can be used in spatial aggregation to provide land cover and land-cover change data at coarser resolutions.

Present schedules call for the first Land Cover Parameter to be released about 15 months after launch based on at least one year of 32-day composites. During this interim period, a provisional at-launch product will be provided. The first Land-Cover Change Parameter will be released about 27 months after launch, based on a two-year period of observation.

New Team Member John Townshend has proposed two new postlaunch products. One of these focuses on an “alarm” algorithm applied to 250-m data that recognizes land cover change, especially deforestation, using changes in spatial texture. The other provides continuous fields of land cover properties, such as proportion in woody cover, using linear mixture modeling.

1.3 Document Scope

The remainder of this document is organized into four broad sections. Section 2 will discuss the rationale for the development of the Land Cover Product, provide a historical context and background information, describe the system of land cover units, and discuss MODIS in terms of its ability to facilitate and improve capabilities for characterizing land cover and land-cover change on a global scale. The objective of section 3 is to define the overall structure of the algorithms; present the mathematics and practical descriptions of the primary algorithm components; present the use of training sites in calibration and validation; discuss sources of error and uncertainty; address practical issues that are likely to arise; discuss preprocessing considerations, computing needs, and reliance on other MODIS activities; detail the calibration and validation phase of the algorithms; and discuss postlaunch product validation. Section 4 addresses potential constraints and limitations. The concluding section 5, contributed by new MODIS Team Member John Townshend, describes the 250-m Land-Cover Change Parameter and additional planned postlaunch products in greater detail.

1.4 Applicable Documents and Publications

The following publications on land cover issues have been supported, in full or in part, by MODIS funding to Boston University.

In Press:

Abuelgasim, A., S. Gopal, J. Irons and A. Strahler. 1996. Classification of ASAS multiangle and multispectral measurements using artificial neural networks. *Remote Sensing of Environment,* Vol. 58: in press.

Moody, A., S. Gopal, and A.H. Strahler.1996. Artificial neural network response to mixed pixels in coarse-resolution satellite data. *Remote Sensing of Environment,* Vol. 58: in press.

Moody, A. and C.E. Woodcock. 1996, Calibration-based models for correction of area estimates derived from coarse resolution land-cover data. *Remote Sensing of Environment.*, Vol. 58: in press.

Strahler, A.H. 1996. Vegetation canopy reflectance modeling—recent developments and remote sensing perspectives. *Remote Sensing Reviews*: in press.

Woodcock, C.E., J.B. Collins and D.L.B. Jupp. 1996. Scaling remote sensing models. in *Scaling Up*, P. Van Gardingen, G. Foody and P. Curran (eds.), Society of Experimental Biology, Cambridge University Press, in press.

Published:

Abuelgasim, A.A., Gopal, S. and A.H. Strahler. 1996. Retrieval of canopy structural parameters from multiangle observations using an artificial neural network. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Lincoln, Nebraska, May 27-31, Vol. 3, pp. 1426-1428.

Borak, J., P. Fisher, A.H.Strahler and A. Moody. 1995. Local-scale evaluation of a technique for land-cover classification based on composited NDVI data, *Proc. 61st Ann. Convention Amer. Soc. Photogrammetry and Remote Sensing,* Charlotte, NC, Feb. 27-March 2, 1995, vol. 3, pp. 796-805.

Borak, J.S. and A.H. Strahler. 1996. Feature selection using decision trees -- an application for the MODIS land cover algorithm. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Lincoln, Nebraska, May 27-31, Vol. 1, pp. 243-245.

Brodley, C.E. and M.A. Friedl.1996. Improving Automated Land Cover Mapping by Identifying and Eliminating Mislabeled Observations from Training Data. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Vol. 2, pp. 1379-1381

Brodley, C.E., M.A. Friedl and A. H. Strahler 1996. New Approaches to Classification in Remote Sensing Using Homogeneous and Hybrid Decision Trees to Map Land Cover., *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium*, Lincoln, Nebraska, May 27-31, Vol. 1, pp. 532-534

Carpenter, G.A., M.N. Gjaja, S. Gopal and C.E. Woodcock. 1996. ART neural networks for remote sensing: Vegetation classification for Landsat TM and terrain data. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Lincoln, Nebraska, May 27-31, Vol. 1, pp. 529-531.

Gopal, S. and M. Fischer. 1996. A comparison of three neural network classifiers for remote sensing classification. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium*, Lincoln, Nebraska, May 27-31, Vol.. 1, pp. 87-789.

Gopal, S., C. Woodcock and A.H. Strahler. 1996. Fuzzy ARTMAP classification of global land cover from AVHRR data set. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Lincoln, Nebraska, May 27-31, Vol. 1, pp. 538-540.

Lambin, E. F and A. H. Strahler. 1994. Indicators of land-cover change for change-vector analysis in multitemporal space at coarse spatial scales. *International Journal of Remote Sensing* 15: 2099-2119.

Lambin, E. F. and A. H. Strahler. 1994. Change-vector analysis: A tool to detect and categorize land-cover change processes using high temporal-resolution satellite data. Remote Sensing of Environment. 8: 231-244. Moody, A. and A.H. Strahler. 1994. Characteristics of composited AVHRR data and problems in their classification. *International Journal of Remote Sensing* 15(17): 3473-3491.

Moody, A. and A.H. Strahler. 1994. Characteristics of composited AVHRR data and problems in their classification. *International Journal of Remote Sensing* 15(17): 3473-3491.

Moody, A. and C. E. Woodcock. 1994. Scale-dependent errors in the estimation of land-cover proportions: Implications for global land-cover datasets. *Photogrammetric Engineering and Remote Sensing* 60: 585-594.

Moody, A. and C.E. Woodcock. 1995. Calibration-based methods for correcting coarse resolution estimates of land-cover proportions. *Proc. 61st Ann. Convention Amer. Soc. Photogrammetry and Remote Sensing,* Charlotte, NC, Feb. 27-March 2, 1995, vol. 3, pp. 684-693.

Moody, A. and C. E. Woodcock. 1995. The influence of scale and spatial characteristics of landscapes on land-cover mapping using remote sensing. *Landscape Ecology* 10: 363-379.

Muchoney, D.M. 1996. The Role of SPOT Data in Global Ecosystem and Biodiversity Monitoring. Proceeding*10th SPOT Conference,* Paris, April 1996.

Muchoney, D.M., J. Borak and A. Strahler. 1996. Global Landcover Classification Training and Validation Issues and Requirements. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Lincoln, Nebraska, May 27-31, Vol. 1, pp. 233-235.

Running, S.W., C.O. Justice, V.V. Salomonson, D. Hall, J. Barker, Y.J. Kaufman, A.H. Strahler, A.R. Huete, J.-P. Muller, V. Vanderbilt, Z.M. Wan, P. Teillet and D. Carneggie. 1994b. Terrestrial remote sensing science and algorithms planned for EOS/MODIS. *International Journal of Remote Sensing* 15(17): 3587-3620.

Strahler, A.H. 1993. *Report of LandCcover Products Workshop: MODIS Land Team*. Flathead Lake, Montana, Sept. 21-23, 1992.

Strahler, A.H.1996. An overview of MODIS products for land applications. *Proc. 8th Australasian Remote Sens. Conf.,* Canberra, Australia, March 25-29, CD-ROM, 8 pp.

Strahler, A.H.1996. AVHRR applications in support of MODIS product development, *Proc. 8th Australasian Remote Sens. Conf.,* Canberra, Australia, March 25-29, CD-ROM, 4 pp.

Strahler, A.H., A. Moody, and E. Lambin. 1995, Land cover and land-cover change from MODIS, *Proc. 1995 International Geoscience and Remote Sensing Symposium,* Florence, Italy, July 10-14, 1995, vol. 2, pp. 1535-1537.

2. Overview and Background Information

2.1 Experimental Objective

Land cover, and human and natural alteration of land cover, play a major role in global-scale patterns of climate and biogeochemistry of the earth system. Although the oceans are the major driving force for the Earth's physical climatology, the land surface has considerable control on the planet's biogeochemical cycles, which in turn significantly influence the climate system through the radiative properties of greenhouse gases and reactive species. Further, variations in topography, albedo, vegetation cover, and other physical characteristics of the land surface generate variations of weather and climate by forcing atmospheric circulation patterns that are driven by surface-atmosphere matter and energy fluxes and the momentum of the earth's rotation.

In this context, an important application of accurate global land-cover information is the inference of parameters that influence biophysical processes and energy exchanges between the atmosphere and the land surface as required by regional and global-scale climate and ecosystem process models (Townshend *et al.*, 1991). Examples of such parameters for climate modeling include leaf area index (LAI), roughness length, surface resistance to evapotranspiration, canopy greenness fraction, vegetation density, root distribution, and fraction of photosynthetically-active radiation absorbed (FPAR) (Sellers, 1991). These serve as input variables that control surface energy and mass balances. Examples of ecosystem process model parameters for which land cover type may serve as a surrogate include leaf photosynthetic capacity, canopy conductance, type of photosynthetic system, and maximum photosynthetic rate (Running and Coughlan, 1988).

Most of these inferences are based on the structural character of the vegetation cover, which is sensible to remote sensing. The objective of the Land Cover Parameter is to identify a suite of land cover types amenable to such parameterization by exploiting the spectral, temporal, spatial, and directional information content of MODIS data. The objective of the Land-Cover Change Product is to detect and quantify the changes in land covers and the natural and anthropomorphic processes that bring them about so that global and regional models may be projected forward through changes in their driving surface parameters.

2.2 Historical Perspective

2.2.1 Global-Scale Land Cover Data

Land-cover datasets currently used for parameterization of global climate models are typically derived from a wide range of preexisting maps and atlases (Olson and Watts, 1982; Matthews, 1983; Wilson and Henderson-Sellers, 1985). This approach has several limitations. First, the reference sources may themselves present a range of different dates, spatial scales, and classification schemes. Confusion regarding the mapping of the reference class units to the classification system and scale used in the land-surface dataset may then lead to errors in the final product. For example, floristic and climatically-based classifications, while not inherently compatible, may need to be combined and reclassified to generate physiognomic cover types for a land-cover compilation (Townshend *et al.*, 1991). Second, the resulting datasets are fundamentally static, and can be assumed to perpetuate errors existing in the sources from which they were derived. Third, some datasets are maps of potential or climax vegetation, which is inferred from climatic variables such as temperature and precipitation rather than the true vegetation type.

Townshend *et al.* (1991) compared the areal statistics and spatial distributions of a large number of land cover datasets. They documented considerable disagreement in the relative percentages of basic land-cover types among these products. Moreover, even when there was general agreement in the relative area covered by given vegetation types, the spatial distribution of these units differed. While such datasets have obvious limitations, they represent the state of the science for driving large scale process models, and have been designed specifically for this purpose.

Many researchers have attempted to produce regional-scale land cover datasets using coarse spatial-resolution, high temporal-frequency data from the AVHRR instrument aboard the NOAA series of meteorological satellites. Almost without exception, these efforts have involved the conversion of AVHRR bands 1 and 2 to NDVI values. The NDVI, being simply the ratio of the difference in near-infrared and red reflectance normalized by their sum. A registered time series of NDVI images is then composited so that, for every pixel location, the maximum NDVI value encountered throughout the compositing period is output. The compositing procedure tends to select against measurements that are strongly influenced by atmospheric and aerosol scattering. These measurements have reduced NDVI values due to differential scattering effects in red and near-infrared bands. Cloud-contaminated measurements also produce lower NDVI values, as clouds reflect strongly in both the red and near-infrared wave bands. The compositing of NDVI values further reduces the variability associated with changing view and illumination geometry (Holben, 1986), although measurements near the forward-scattering direction tend to have higher NDVI values and will thus be preferentially selected. Compositing periods are chosen based on a trade-off between the expected frequency of changes in vegetation and the minimum length of time necessary to produce cloud-free images.

The NDVI generally quantifies the biophysical activity of the land surface and, as such, does not provide land cover type directly. However, a time series of NDVI values can separate different land cover types based on their phenology, or seasonal signals (*e.*g. Lenney *et al*., 1996). Reed *et al.* (1994) and DeFries *et al.* (1995) have developed and used multitemporal phenological metrics to derive land cover classifications from AVHRR data. Lambin and Ehrlich (1996a and 1996b) have found that using a time series of the ratio of surface temperature to NDVI provides a more stable classification than NDVI alone, primarily by isolating interannual climatological variability.

Townshend, *et al.* (1987) performed supervised classifications on composited NDVI GAC (Global Area Coverage) data for South America. While they did not validate their results with test data, they found that accuracy for the training sites improved substantially with the increase in the number of images included in the time series. Koomanoff (1989) used annually-integrated NDVI values to generate a global vegetation map using NOAA's Global Vegetation Index product (GVI).This work represents nine vegetation types and does not rely on the seasonality of the NDVI. Lloyd (1990) employed a binary classifier based on summary indices derived from a time series of NDVI data. These phytophenological variables included the date of the maximum photosynthetic activity, the length of the growing season, and the mean daily NDVI value. The variables were fed through a binary decision tree classifier that stratified pixels based first on the date of the maximum NDVI, then the length of the growing season, and finally on the mean daily NDVI. Although the author offers no validation, the approach is intuitively appealing and provides a framework that could expand to include stratification based on different factors at different scales.

A coarse-resolution global land surface parameter database has recently been released on five compact disks as an activity of the International Satellite Land Surface Climatology Project (ISCLSCP, Meeson *et. al*. 1995). The database includes land cover classes, absorbed fraction of photosynthetically-active radiation (FPAR), leaf area index (LAI), roughness length, and canopy greenness fraction, along with data on global meteorology, soils, and hydrology. The spatial scale of the database is 1° by 1°. Variables such as FPAR, LAI and canopy greenness fraction are derived from 8-km composited AVHRR NDVI data. The land cover classification is based on a spatial aggregation of the 8-km data to one degree followed by supervised classification of the temporal patterns in NDVI (DeFries and Townshend, 1994).

Loveland *et al.* (1991, 1995) used one year of monthly composited AVHRR-LAC data to generate an unsupervised classification of land cover types for the conterminous United States. The resulting clusters were further stratified based on ancillary environmental data such as elevation and ecoregion. Class labels were assigned based on the temporal curves of the clusters as well as a large number of ancillary sources. While obviously limited by the quality of the composited NDVI data, as well as the accuracy of the ancillary sources, this dataset represents the most convincing large area classification of AVHRR data at 1-km spatial resolution to date.

At present, Loveland's effort continues under the auspices of the IGBP-DIS (International Geosphere Biosphere Programme-Data and Information System), based on a global database of 1-km AVHRR observations received during the period April 1992 through September 1993 (Belward and Loveland, 1995; Belward, 1996). These have been assembled and 10-day composited at the EROS Data Center (EDC). The 10-day composites have been further composited into monthly composites, thus providing a multitemporal database for land cover classification using an unsupervised clustering and labeling approach. The product is being developed in continental segments; as of November 1996, North America, South America, and Africa had been completed, with the remaining continents to be completed by June, 1997.

2.2.2 Land Cover Classification Using Neural Networks

In the recent literature, several new approaches to pattern recognition have been brought to bear in the problem of classification of remotely sensed data. These include artificial neural networks, decision-tree classifiers, and computer vision algorithms. Thus far, the most promising classification techniques seem to be based on neural networks. Accordingly, we have identified the neural network approach as the candidate algorithm for the Land Cover Parameter. The following discussion reviews the literature on land cover classification using neural networks.

The classification of remotely sensed datasets using artificial neural networks first appeared in the remote sensing literature about five years ago. Since then, examples and applications have become increasingly common. Remotely-sensed datasets processed by neural network-based classifiers have included images acquired by the Landsat Multispectral Scanner (MSS) (Lee *et al.*, 1990; Benediktsson *et al.*, 1990); Landsat TM (McClellan *et al.*, 1989; Hepner *et al.*, 1990; Kiang, 1992; Bischof *et al.*, 1992; Heermann and Khazenie, 1992; Salu *et al.*, 1993; Yoshida and Omatu, 1994; Fitzgerald and Lees, 1994;); synthetic aperture radar (Decatur, 1989; Hara et. al, 1994); SPOT (Kanellopoulos *et al.*, 1992; Dreyer, 1993; Tzeng *et al.*, 1994); AVHRR (Key *et al.*, 1989; Gopal *et al.*, 1994); and aircraft scanner data (Benediktsson *et al.*, 1993). A number of these studies have also included ancillary data (*e.g.*, topography, as in Benediktsson *et* *al.*, 1990, 1993, or texture, as in Bischof *et al.*, 1992). Many studies have been directed toward recognition of land cover classes, which have ranged from broad life-form categories (*e.g.*, Hepner *et al.*, 1990) to floristic classes (Fitzgerald and Lees, 1994). Most use a supervised approach, but unsupervised classification using self-organizing neural networks has also been attempted (*e.g.*, Hara et. al, 1994). In nearly all cases, the neural network classifiers have proved superior to conventional classifiers, often recording overall accuracy improvements in the range of 10-20 percent.

As the number of successful applications of neural network classification increases, it seems increasingly clear that neural network-based classification is superior to conventional approaches for remote sensing. The reasons include: (1) neural network classifiers are distribution-free and can detect and exploit nonlinear data patterns; (2) neural network classification algorithms can easily accommodate ancillary data; (3) neural network classifiers are typically more accurate than conventional classifiers; and (4) neural network architectures are quite flexible and can be easily modified to optimize performance.

Neural networks can also yield important information about the decision structure contained within them. For example, Key at al. (1989) examined the weights and interconnections within the layers of a trained neural network to provide insight into the way that the network makes decisions. More recently, Gopal and Woodcock (1995) successfully used principal components analysis (PCA) to interpret the structure of weights in a fully-trained network for a change detection study.

Much of the neural net classification work in remote sensing has used multilayer feed-forward networks that are trained using the backpropagation algorithm, a recursive learning procedure using a gradient descent search. However, this training procedure is sensitive to the choice of initial network parameters and to overfitting (Gopal *et al.*, 1994). The use of nets utilizing adaptive resonance theory (ART) (Carpenter, 1989) has overcome these problems. Networks organized on the ART principle are stable as learning proceeds, while at the same time they are plastic enough to learn new patterns.

Some recent studies by Gopal *et al.* (1994) highlight the utility of fuzzy ART neural network architecture and its application to the MODIS land cover classification problem. Fuzzy ARTMAP automatically constructs a minimal number of recognition categories to meet accuracy criteria. A voting strategy improves prediction by training the system several times on different orderings of an input set. ARTMAP dynamics are fast, stable, and scalable, overcoming common limitations of back propagation.

In the first study, Gopal *et al.* (1994) classified one year of monthly composited AVHRR LAC NDVI data from west Africa at 1.1-km spatial resolution into six broad life-form classes using a fuzzy ARTMAP neural net classifier (Carpenter *et al.*, 1991). The authors report per-class accuracies ranging between 66 and 98 percent, with an average accuracy of 83 percent. The latter result compares very favorably with a more typical feed-forward architecture trained by back propagation that achieved an average accuracy of only 61 percent. Note that feed-forward back-propagation networks have generally been observed to outperform conventional supervised classifiers, such as maximum likelihood.

Using the annual sequence of composited Normalized Vegetation Index (NDVI) values in an AVHRR data set, DeFries and Townshend (1994) distinguished eleven global cover types and classified global land cover with a maximum likelihood classifier. Gopal *et al.* (1996) used fuzzy ARTMAP to classify the same data set. Classification results were analyzed and compared with those obtained from DeFries and Townshend (1994). First, classification accuracy on the training data set is 100 percent compared with the corresponding accuracy of 86.8 percent. Second, when fuzzy ARTMAP is trained using 80 percent of data set and tested on the remaining (unseen) 20 percent data set, classification accuracy is more than 80 percent. Third, classification results vary with and without latitude as an input variable similar to those of DeFries and Townshend.

In another study, Gopal and Fischer (1996) evaluated the performance of Multi-Layer Perceptron (MLP), Radial Basis Function, and Fuzzy ARTMAP networks using a Landsat-5 TM scene of the northern section of the city of Vienna, Austria. Classification accuracies obtained from the neural network classifiers are compared with a benchmark, the maximum likelihood classifier. In addition to the evaluation of classification accuracy, the neural networks are analyzed for their generalization capability and stability of results. Best overall results (in terms of accuracy and convergence time) are obtained using fuzzy ARTMAP followed by MLP (with weight elimination). Their classification error rates on the training data set are zero and 7.87 percent respectively; classification error on the testing data set are 10.24 percent and less than 2 percent. Simulation results serve to illustrate the properties of various classifiers in general, as well as the stability of the results with respect to various critical control parameters, initial parameter conditions, training time, and different training and testing data sets.

Carpenter *et al.* (1996) developed a new methodology for automatic mapping from Landsat Thematic Mapper (TM) and terrain data, based on the fuzzy ARTMAP neural network. System capabilities were tested on a challenging remote sensing classification problem, using spectral and terrain features for vegetation classification in the Cleveland National Forest. After training at the pixel level, system capabilities were tested at the stand level, using sites not seen during training. Results were compared to those of maximum likelihood classifiers, as well as back propagation neural networks and K Nearest Neighbor algorithms. Best results were obtained using a hybrid system based on a convex combination of fuzzy ARTMAP and maximum likelihood predictions. A voting scheme was used to assign confidence estimates to competing predictions.

In another recent study with implications for MODIS land cover, Moody *et al.* (1996) showed that neural net output vectors need not be interpreted categorically, but under some circumstances can be used as fuzzy predictors of class membership (Birdle, 1990). Moody *et al.* demonstrate that where pixels are mixed, neural net outputs can detect such mixtures, a result that leads the way for both automatic categorization of mosaic classes of mixed covers and the generation of confidence flags when single-category outputs are assigned.

2.2.3 Directional Information in Land Cover Classification

Although the application of spectral, temporal, and spatial information in classification of remotely-sensed data has long been established in the literature, it is only recently that the importance of directional information has been established. In a study utilizing directional aircraft scanner imagery of an agricultural test site, Barnsley (Barnsley *et al.*, 1990; Barnsley, 1995) noted that by using a principal components transformation and adding three view angles to two bands of spectral data, classification accuracies increased by approximately 20 percent. Abuelgasim and Gopal (1994; Abuelgasim *et al.*, 1995) used a hybrid unsupervised-supervised neural network classifier to distinguish five broad land cover classes in Minnesota subboreal forest from directional reflectances imaged by NASA's ASAS (Advanced Silicon Array Spectrometer) aircraft instrument (Irons *et al.*, 1991). Using seven directional images acquired in a single near-infrared band, the authors obtained 89 percent accuracy using the hybrid classifier, as compared to 85 percent for a conventional feed-forward back-propagated neural net and 61 percent for a maximum likelihood classifier.

2.2.4 Land-Cover Change

Global assessment of the changes in physical characteristics of the terrestrial surface cover is a fundamental input for models of global climate and terrestrial hydrology. While some changes in land cover, such as long-term changes in climate due to astronomical causes, or shorter-term vegetation successions produced by geomorphic processes are caused by natural processes, human activity increasingly modifies the land surface cover. These modifications arise through direct actions, such as deforestation, agricultural activities and urbanization, or indirectly, through human-induced climatic change. The importance of mapping, quantifying, and monitoring changes in the physical characteristics of land cover has been widely recognized in the scientific community as a key element in the study of global change (*e.g.* IGBP, 1994a; Henderson-Sellers and Pitman, 1992; Nemani and Running, 1995, 1996).

Digital change detection is the process of determining and/or describing change based on coregistered, multitemporal remotely sensed data. The two principal approaches to change detection are 1) post-classification techniques, where independent classifications are compared and 2) pre-classification or merged data techniques where simultaneous analysis of multitemporal data occurs (Malila, 1980; Muchoney and Haack, 1994). Post-classification techniques have significant limitations. The comparison of land cover classifications for different dates does not allow the detection of subtle changes within a land cover class. Also, the change map product of two classifications exhibits accuracies similar to the product of multiplying the accuracies of each individual classification (Stowe *et al.*, 1980). Merged data techniques include image differencing/ratioing, change vector analysis, spectral-temporal (layered-temporal) change classification, regression techniques and principal components analysis.

The Land-Cover Change Parameter employs the pre-classification or merged data approach. Rather than analyzing isolated dates from two separate time periods, it is based on a comparison of the temporal development curve, or time-trajectory, for successive years of indicators. The indicators, derived from remotely sensed data, include such variables as vegetation indexes, surface temperature, or spatial structure (Lambin and Strahler, 1994a) and are provided by the 32-day composited database assembled for land cover classification.

2.2.4.1 Change Vector Analysis

The primary change detection technique for the 1-km Land-Cover Change Parameter is change vector analysis (Lambin and Strahler, 1994b). In this technique, each annual multitemporal set of indicator values is taken as a point in multitemporal space, and points from successive years are connected by a change vector, also in multitemporal space. The direction of the change vector quantifies the change process, while the magnitude of the change vector quantifies the amount of change (Lambin and Strahler, 1994b). Change vectors applied to different indicators reveal different change processes or different aspects of change processes (Lambin and Strahler, 1994a).

The application of change vectors in remote sensing was first described by Malila (1980) and by Colwell and Weber (1981), although the change vectors in these studies were multispectral rather than multitemporal. Michalek *et al.* (1993) applied multispectral change vectors to the monitoring of coastal environments. Change vector analysis in the temporal domain lends itself to AVHRR applications, since this instrument provides data with a high temporal frequency. Prior studies of change using AVHRR have used annual integrated NDVI, or isolated dates of NDVI or untransformed data. Examples are the studies of change in the Sahel of Tucker *et al.* (1986, 1991), Hellden and Eklundh (1988) and Hellden (1991); or studies of large-scale tropical deforestation by Tucker *et al.* (1984), Nelson and Holben (1986), Woodwell *et. al*. (1987), and Malingreau *et al.* (1989). While the procedures used in these studies are appropriate to detect abrupt land-cover changes such as forest clearing, biomass burning, or the impact of a severe drought, the detection of more subtle forms of change, such as those associated with climate change or with slow rates of land degradation, requires a more sophisticated approach such as change vector analysis.

Change vector analysis has been explored in several studies that have been partially supported by, or coordinated with, the MODIS 1-km land cover effort. These have focused on multitemporal datasets of Africa at LAC and GAC resolutions made available by J.-P. Malingreau (Joint Research Center, Ispra, Italy). In a study of two years of LAC NDVI data in the Sahel region, change vectors were calculated using monthly maximum-value composites, then subjected to principal components analysis (Lambin and Strahler, 1994b). The components were related to: 1) the timing of the start of the growing season; 2) vegetation senescence rates in savannas during the onset of the dry season; 3) vegetation senescence rates in herbaceous covers during the onset of the dry season; and 4) differences produced by haze and cloud contamination (Lambin and Strahler, 1994a). The analysis clearly demonstrated the ability of the technique to detect subtle variations in regional phenology, thus providing a basis for separating natural temporal variability from more permanent changes induced by human activity.

This analysis was extended to compare maximum-value composites of NDVI with maximum-value composites of surface temperature and spatial structure (Lambin and Strahler, 1994a). Spatial structure was quantified by calculating the standard deviation of NDVI values within an adaptive three-by-three pixel window. In the adaptive window procedure, the standard deviation is computed for each of nine three-by-three windows to which a pixel belongs, and the minimum value is selected (Woodcock and Ryherd, 1989).

In this way, the texture measure is not inflated artificially by the contrast between land cover boundaries.

The analysis showed these indicators to have a low degree of redundancy. NDVI change vectors are driven by seasonal changes in rates of vegetation activity; surface temperature change vectors are driven on a shorter time scale by rainfall events, especially in the drier environments; and NDVI texture displays a seasonal variability that must be taken into account when assessing long-term change. Later studies of change in NDVI texture in the same region confirm the diagnostic nature of the temporal pattern of spatial heterogeneity (Lambin, 1995).

The change vector technique is the most mature of the available techniques suited to the 1-km Land-Cover Change Product, and thus will be the basis for initial 1-km Land-Cover Change Product provided early in the postlaunch period. Other approaches will be explored simultaneously, especially at our intensive study sites. These techniques include neural network-based change detection, principal components analysis and Gramm-Schmidt orthogonalization.

2.2.4.2 Neural Network-Based Change Detection

Artificial neural networks have only recently been applied to change detection. Gopal *et al.* (1996) have been evaluating the use of the fuzzy ARTMAP neural network for land cover classification and change detection at global scales using multitemporal AVHRR NDVI. These techniques have also been applied locally in a supervised approach to detect changes in forest cover attributes over time (Gopal and Woodcock, 1996).

The artificial neural network used in a supervised approach to develop the Land Cover Parameter, by its nature, includes a change detection component. As new data are presented to the Fuzzy ARTMAP neural network, the input either matches an existing category, or a new category must be created. If a new data presentation does not match an existing category then it will be necessary to determine whether the new data represent a fundamentally new condition (change) or whether the vigilance parameter needs to be relaxed so that an existing category (presumably a pixel’s previous category) can now accommodate the new input. This feature is especially useful for flagging change in the training process, when training data may be from an earlier period.

Another feature of neural network classifiers is that probability of membership by class can be evaluated on a pixel-by-pixel basis to track the probability that a specific pixel belongs to a certain class. When a specified threshold has been crossed, that pixel has moved into a new class and a categorical change has occurred. For the 1-km Land-Cover Change Parameter, a direct supervised approach to change detection using neural networks (Gopal and Woodcock, 1996) will also be investigated for targeting specific types of change in specific areas.

2.2.4.3 Principal Components Analysis

Principal components analysis, or the related Karhunen-Loeve (K-L) transformation (Duvernoy and Leger, 1980), is a multivariate statistical technique in which data axes are rotated into principal axes, or components, that maximize data variance. The original data are then transformed to the new principal axes, or components. In this manner, correlated data sets can be represented by a smaller number of axes, while maintaining most of the variation of the original data. PCA has been widely applied to detect and to isolate and determine the nature of changes in the remote sensing signal over time (Byrne *et al.*, 1980; Muchoney and Haack, 1994). For the 1-km Land-Cover Change Parameter, PCA will serve two functions. First, it will be used as part of the QA procedure, as a means of evaluating the nature of variance in the time series for anomalies and artifacts due to sensor characteristics and data processing. Second, PCA will be used at the intensive studies sites for land cover change to evaluate the performance of the change vector and neural network techniques.

2.2.4.4 Gramm-Schmidt Transformation

The Gramm-Schmidt orthogonalization change detection technique (Collins and Woodcock, 1994) will be employed in the QA process to evaluate the global change detection algorithm (change vector). It may also be used to discriminate specific types of change or region-specific change, and to describe the nature of change as determined by change vector analysis. Gramm-Schmidt is a more physically-based, less-empirical approach to rotational transformation than PCA. Gramm-Schmidt change detection is a modification of the technique that was initially used to derive coefficients of the tasseled-cap transformation for single-date imagery (Kauth and Thomas, 1976) to accommodate multitemporal data. In this case, the coordinate scores of rotated multitemporal pixel vectors directly represent inter-date change.

2.3 Land Cover Units

The primary objective of the land cover parameter is to facilitate the inference of biophysical information from land cover for use in regional and global modeling studies. Thus, the specific classification units of land cover need not only to be discernible with high accuracy from remotely-sensed and ancillary data, but also need to be directly related to physical characteristics of the surface and primarily to surface vegetation. A set of 17 such global land cover classes has been developed by the IGBP-DIS in conjunction with the IGBP Core Projects specifically for this purpose (Belward, 1996). They are to be applied to a classification of the global 1-km composited AVHRR LAC NDVI database assembled at EROS Data Center that is now nearing completion (Belward and Loveland 1995, Belward, 1996). Since the IGBP system of units was developed for a global land cover product at a similar resolution, 1 km for a similar purpose (biophysical parameterization for modeling) we plan to use the IGBP classification for the MODIS Land Cover Product.

Table 1 provides a list of the IGBP land cover units with accompanying descriptions. The list includes eleven classes of natural vegetation, three classes of developed and mosaic lands, and three classes of nonvegetated lands. The natural vegetation units distinguish evergreen and deciduous, broadleaf and needleleaf forests, where one of each pair of attributes dominates; mixed forests, where mixtures occur; closed shrublands and open shrublands; savannas and woody savannas; grasslands; and permanent wetlands of large areal extent. The three classes of developed and mosaic lands distinguish among croplands, urban and built-up lands, and cropland/natural vegetation mosaics. Classes of non-vegetated land cover units include snow and ice; barren land; and water bodies.

Note that the IGBP classes can be relabeled (cross-walked) to provide compatibility with current and future systems used by the modeling community. Table 2 provides an example in which the IGBP classes are translated to three other schemes. SiB2 (Sellers, *et. al*, 1994) is a surface-atmosphere interaction model for use in GCMs; the classification of Running and Nemani (Nemani and Running, 1996; Running *et. al*., 1994a; Running *et. al*., 1995) is intended primarily for parameterization of global carbon and nutrient cycle models; and Myneni's classification is used in radiative transfer modeling for the MODIS and MISR LAI/FPAR products. For nearly all classes in these schemes, there is a direct mapping of one or more IGBP classes to their equivalents. A problem arises where some classes have no equivalents. For example, wetlands and urbanized areas do not appear in these three schemes, presumably because they are not recognized in the models receiving the land cover type input. However, these classes may be quite important for trace gas emission models, for example. The IGBP strategy was to recognize those classes that would be most useful across all the modeling disciplines of the IGBP, in effect requesting modelers to consider all relevant classes with significant areal extent on the earth's land surface.

2.3.1 Biophysical Parameterization

Taken together, the sets of natural vegetation and developed lands units can be used to differentiate several fundamental distinctions among cover types that are essential for ecological process modeling. One of these is annual vs. perennial habit, distinguished by whether or not the vegetation retains perennial or annual aboveground biomass. This attribute separates vegetation with permanent respiring biomass (forests and woody-stemmed shrubs) from annual crops and grasses that go through nongrowing season periods as seeds or below-ground structures only. The annual-perennial distinction allows inference of several critical physiological attributes of plants. For example, in a global synthesis of plant gas exchange rates, Korner (1993) found on average that annual plants maintained a 50 percent higher leaf photosynthetic capacity than perennial plants. Biomass permanence, as it relates to plant height, also is the major vegetation determinant of the surface roughness length parameter that climate models require for energy and momentum transfer equations.

Another fundamental attribute that is distinguishable using this set of units is leaf longevity, which distinguishes between evergreen and deciduous plant covers. This attribute is a critical variable in carbon cycle dynamics of vegetation, and is important for seasonal albedo and energy transfer characteristics of the land surface. The leaf longevity class defines whether a plant must completely re-grow its entire canopy each year, or merely a portion of it, with direct consequences to ecosystem carbon partitioning, leaf litterfall dynamics and soil carbon pools. Reich *et al.* (1992) suggests that canopy conductance and maximum photosynthetic rate are inversely proportional to leaf longevity. Hence, certain global attributes of canopy gas exchange capacity may be inferred based on a leaf longevity criterion.

Table 1. IGBP Land Cover Units	
Natural Vegetation	
Evergreen Needleleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Almost all trees remain green all year. Canopy is never without green foliage.
Evergreen Broadleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Almost all trees and shrubs remain green year round. Canopy is never without green foliage.
Deciduous Needleleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Consists of seasonal needleleaf tree communities with an annual cycle of leaf-on and leaf-off periods.
Deciduous Broadleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Consists of broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods.
Mixed Forests	Lands dominated by trees with a percent cover >60% and height exceeding 2 meters. Consists of tree communities with interspersed mixtures or mosaics of the other four forest types. None of the forest types exceeds 60% of landscape.
Closed Shrublands	Lands with woody vegetation less than 2 meters tall and with shrub canopy cover >60%. The shrub foliage can be either evergreen or deciduous.
Open Shrublands	Lands with woody vegetation less than 2 meters tall and with shrub canopy cover between 10-60%. The shrub foliage can be either evergreen or deciduous.
Woody Savannas	Lands with herbaceous and other understory systems, and with forest canopy cover between 30-60%. The forest cover height exceeds 2 meters.
Savannas	Lands with herbaceous and other understory systems, and with forest canopy cover between 10-30%. The forest cover height exceeds 2 meters.
Grasslands	Lands with herbaceous types of cover. Tree and shrub cover is less than 10%.
Permanent Wetlands	Lands with a permanent mixture of water and herbaceous or woody vegetation. The vegetation can be present in either salt, brackish, or fresh water.
Developed and Mosaic Lands	
Croplands	Lands covered with temporary crops followed by harvest and a bare soil period (*e.g.*, single and multiple cropping systems). Note that perennial woody crops will be classified as the appropriate forest or shrub land cover type.
Urban and Built-Up Lands	Land covered by buildings and other man-made structures.
Cropland/Natural Vegetation Mosaics	Lands with a mosaic of croplands, forests, shrubland, and grasslands in which no one component comprises more than 60% of the landscape.
Non-Vegetated Lands	
Snow and Ice	Lands under snow/ice cover throughout the year.
Barren	Lands with exposed soil, sand, rocks, or snow and never has more than 10% vegetated cover during any time of the year.
Water Bodies	Oceans, seas, lakes, reservoirs, and rivers. Can be either fresh or salt-water bodies.

A third vegetation attribute recognizable within the IGBP units is the leaf type or shape of the dominant vegetation cover. Three leaf shapes are distinguished among the various categories: needleleaf, broadleaf, and graminoid (grasses). This attribute also correlates well with key ecological parameters for biogeochemical modeling. Running and Hunt (1993) defined maximum leaf area index values of 10, 6 and 3, and maximum

Table 2. Classification System Comparisons

Classification

IGBP	SiB2 Biome	Running & Nemani	Myneni
Objective			
General Model Parameterization	Surface Interactions in GCMs	Carbon and Nutrient Cycling Models	Radiative Transfer for LAI, FPAR
Class Breakdown			
Evergreen Needleleaf Forests	Needleleaf-Evergreen Trees (4)	Evergreen Needleleaf	Needle Forests
Deciduous Needleleaf Forests	Needleleaf-Deciduous Trees (5)	Deciduous Needleleaf	Needle Forests
Evergreen Broadleaf Forests	Broadleaf-Evergreen Trees (1)	Evergreen Broadleaf	Leaf Forests
Deciduous Broadleaf Forests	Broadleaf-Deciduous Trees (2)	Deciduous Broadleaf	Leaf Forests
Mixed Forests	Broadleaf and Needleleaf Trees (3)		
Woody Savannas	C-4 Grassland (6)	Savannas	Savanna
Savannas	C-4 Grassland (6)	Savannas	Savanna
Grasslands	C-4 Grassland (6)	Grasses	Grasses/Cereal Crops
Closed Shrublands	Dwarf Trees and Shrubs (8)		Shrublands
Open Shrublands	Shrubs with Bare Soil (7)		Shrublands
Croplands	Agriculture or C-3 Grassland (9)		Broadleaf Crops
Cropland/Natural Vegetation Mosaics			
Permanent Wetlands			
Urban and Built-Up Lands			

canopy conductance values of 1.6, 2.5, and 5.0 mm/sec, for needleleaved tree, broadleaved tree, and grass covers, respectively.

Still other intrinsic biophysical parameters may be inferred from these units. For example, Dorman and Sellers (1989) assigned a series of optical properties, physiological properties and physical parameters to a set of vegetation classes, based on those of Mathews (1983) and Kuchler (1983), specifically for global application to the Simple Biosphere Model of land surface-atmosphere interaction (Sellers *et al.*, 1986). In a somewhat different application, Sellers *et al.* (1994) devised an algorithm for determining global FPAR, LAI, and canopy greenness fraction from monthly composited NDVI at 1-degree resolution. Although the method is based on NDVI, their algorithm stratifies NDVI-FPAR-LAI relationships by vegetation cover types, using broad structural classes similar to those of the IGBP classification. Note that FPAR and LAI will be produced from MODIS using separate biogeophysical algorithms in MODIS Product MOD15.

2.3.2 At-Launch Provisional Land Cover Parameter

Because input for the MODIS Land Cover Parameter will not be prepared until a time-trajectory of radiometric measurements spanning at least one year is acquired, there is a need for a provisional, at-launch parameter. This provisional parameter will be used by the Land Team in preparation of several at-launch land products, including LAI/FPAR (MOD15), annual net primary productivity and photosynthesis parameter (MOD16, MOD17), snow cover (MOD10), and BRDF/Albedo (MOD43), to infer various types of biophysical information as required for their investigations.

The likely source of this provisional parameter is the IGBP global land cover data-base described above (Belward and Loveland, 1995; Belward, 1996). This database is being constructed by applying an unsupervised clustering and labeling technique to the global composited NDVI 1-km database assembled at EROS Data Center. The dataset is being prepared in five segments (North America, South America, Africa, Europe, and Asia); at the time of writing of this document, North and South America were complete, and Africa was nearly complete. The project is scheduled to release a global product in June, 1997.

It will be important provide a quality assessment of the at-launch product. Validation data being generated to test the IGBP product could be one such source. Other sources of information will also be used including validation data sets derived from results of the Landsat Pathfinder Global Land Cover Test Sites and Humid Tropical Pathfinder projects. These various products should be used to assess the quality of the IGBP product and other similar products as available.

2.4 Spatial Resolution

The Land Cover Product will be produced at 1-km spatial resolution. This scale is the finest that is practically achievable with the MODIS instrument. Although the MODIS land bands are imaged at 250- and 500-m spatial resolution, these are nominal values for nadir pixels. At the edge of the MODIS swath, pixels grow by a factor of 2 in the along-track direction and by a factor of 5 in the across-track direction. In addition, there will be geolocation error in computing the center of each pixel that is produced by uncertainties in the knowledge of the location of the spacecraft and its orientation. The best current estimates of pointing error are ±85 m along-track and ±153 m across-track (MODIS geolocation workshop, 8/8/96, GSFC). These are three standard-deviation (3 sigma) values projected to nadir, and will increase with look angle in a fashion similar to that of pixel size.

Geolocation errors, together with pixel size growth, give an effective instantaneous field of view (EIFOV) of 656 m in the along-track direction and 960 m in the across-track direction for the composited sequence of measurements input to the 32-day land cover database. This EIFOV does not include optical blurring or scatter within the instrument, and thus is a conservative estimate of effective spatial resolution. Moreover, multidate images will not overlap perfectly, and thus resampling is always necessary. This effect will further smooth the data.

Given these effects, it seems prudent to use a grid cell size of 1-km. Finer resolution information would be redundant and incur significant costs to produce and store.

The 1-km spatial resolution is well-suited to the needs of the global and regional modeling community. In a recent report, the IGBP-DIS Land Cover Working Group noted (IGBP, 1992b):

> "There is an emerging view regarding the appropriate scale for analyzing land cover and land cover conversion. The suitability of 4-8 km GAC data for delineating broad land cover types and phenology has been demonstrated (Malingreau 1986, Malingreau and Tucker 1987). The utility of 8-15 km data for land cover classification and phenology has also been shown by a number of authors (*e.g.*, Justice *et al.*, 1985, and Tucker *et al.* 1985), but it is too coarse for monitoring land cover conversion and reliable detection of land transformation requires resolutions of 1 km or finer (Townshend and Justice 1988). This observation is supported by detailed analyses of tropical deforestation, which suggests that even 1 km data might be too coarse for quantifying the area and rate of deforestation in some regions (C. J. Tucker personal comm.), although a 1 km data set would assist stratified sampling."

These considerations led directly to the development of the IGBP 1-km Land Cover Database described above. In addition, 1-km scale input is required for other 1-km MODIS land products as noted in section 2.3.2.

Note also that at least three recent studies on the effect of spatial resolution on mapping of land cover (Moody and Woodcock, 1994; Mayaux and Lambin, 1995; Pax Lenney and Woodcock, 1996) have shown little difference in proportional errors due to aggregation in the 500 m–1 km range. Rather, it is only at finer resolutions—100 m or less—that proportional errors change. Thus, finer resolution—500 m for example—only serves to increase spatial precision rather than capture a different level of spatial pattern.

2.4.1 Aggregation and Scaling

The effect of scaling on land cover proportions has been explored in the research literature in recent years. Several studies have established that changing the spatial resolution of land cover maps has important effects on the proportion of a landscape occupied by a particular land cover type (Henderson-Sellers *et al.*, 1985; Turner *et al.*, 1989; Moody and Woodcock, 1994; Moody and Woodcock, 1995). In general, the proportions of smaller, more fragmented cover types decrease with aggregation, while those of the larger classes increase. Similar effects were noted by Townshend and Justice (1988), who observed large changes in the proportions of test site images falling within specific NDVI ranges as scenes were progressively degraded to coarser resolutions. These observations conform with more theoretical results obtained by Jupp *et al.* (1988, 1989) and Woodcock and Strahler (1987) on scaling, resolution, and spatial pattern.

The results obtained by these researchers suggest that where different covers differently influence biophysical relationships (*e.g.*, NDVI-FPAR relationships, Sellers *et al.*, 1994), the aggregated behavior of an areal unit will be different from that of a single cover type that dominates it. This effect has been documented in surface energy balance modeling by Henderson-Sellers and Pitman (1992).

There are two further implications of the scaling behavior of land covers for the MODIS product. First, if 1-km land cover classes are to be aggregated to a coarser grid, the product should provide a vector of proportions by classes within coarse grid cells, rather than a label derived from a single dominant class. In this way, users will have sufficient information to treat the area as heterogeneous if desired. This approach will be used in the Land Cover 1/4-degree Climate Modeler's Grid (CMG) Product.

A second implication lies at subpixel scales. Many studies of fine resolution satellite imagery have established the fact that the spatial pattern structure of real landscapes is often finer than 1 km in linear dimension (*e.g.*, Townshend and Justice, 1988; Townshend *et al.*, 1992). These findings suggest that even at 1-km resolution most pixels are mixed. If pixels are mixed and the proportions matter, then it is important to document the way that proportions change with pixel size within ecoregions so that subpixel effects may be accommodated in coarser-scale aggregations (Moody and Woodcock, 1995). Such documentation will require validation studies at test sites (discussed in a later section).

Several approaches to correction of cover-type proportions have been explored in the literature. These have ranged from simple regression methods in which fine-resolution proportions are associated with coarse-resolution spectral variables (Zhu, 1994) to regression-tree models predicting coarse-scale proportions from a suite of fine-scale spatial pattern quantifiers using regression trees (Moody and Woodcock, 1995). In a recent study examining the determination of tropical forest area from AVHRR LAC data, Mayaux and Lambin (1995) provide a two-step procedure in which the Matheron index, which measures the length of boundary per unit area for a class, estimates the slope and intercept for a regression linking fine- and coarse-resolution proportions. They recently extended their work to inverse calibration of cover proportions using several measures of spatial textures, including a simulation of the MODIS scenario of 1-km classification with a standard deviation texture measure derived from a 250-m band (Mayoux and Lambin, 1996). They showed that the 250-m texture was the most effective of the measures available, reducing residuals in observed versus modeled proportions to less than 9 percent.

These studies suggest that areal proportion estimation must be done carefully with appropriate use of fine-scale information, such as 250-m spatial texture. Moreover, proportion estimation is dependent on the spectral characteristics of the classes as well as their spatial structures. For the Land Cover Product, we will provide subpixel proportion estimation information as an ancillary dataset for those regions in which we have sufficient information to compute it. In the post-launch era, we may be able to provide a specific global product with this information.

2.5 Instrument Characteristics

Several problems have been consistently encountered in attempting to process AVHRR data for large-area land-cover discrimination. Although the red and near-infrared channels capture the primary variance in the vegetation signal, other spectral bands that can provide important information in reflective wavelengths are missing. Furthermore, the AVHRR bands are broad and include some atmospheric absorption features that unduly influence the ground signal. Despite the use of the maximum value compositing procedure to screen for clouds, many cloud-covered pixels are still included in composited images. This leads to a misrepresentation of the time trajectory and can cause faulty classification results.

Composited images can also include large numbers of poorly-registered off-nadir pixels, a circumstance that results in a blurred image appearance, reduced image variance, and variable within-scene spatial resolution. Part of the reason for this lies in the instrument's conical scan mirror, which provides rotated and overlapping instantaneous fields of view at the edge of the scan. Other limitations to the processing of AVHRR data include poor spectral and radiometric calibration, poor pointing knowledge, and difficulties in providing accompanying atmospheric correction.

The design of the MODIS instrument alleviates many of these problems. Seven of the MODIS bands in the reflective region have been selected expressly for land applications based on experience with Landsat Thematic Mapper and AVHRR. These bands are positioned to sample the solar spectral curve in wavelength regions that provide specific information about the land surface, while their bandwidths are chosen so as to maximize radiometric precision and avoid atmospheric absorption. This will lead to the production of vegetation indices that are more meaningful and more resistant to atmospheric effects than those produced from AVHRR data, and will also provide for greater utility and interpretability of the individual bands. All of the reflective bands and several combinations of these bands will be potentially useful for discriminating land cover units and monitoring change (Townshend *et al.*, 1991). In addition, thermal channels may also prove to be useful in characterizing and discriminating land cover types (Townshend *et al.*, 1991).

In contrast to AVHRR, MODIS possesses an extensive on-board capability for radiometric and spectral calibration. Calibration subsystems include for shorter wavelengths the spectroradiometric calibration assembly (SRCA), the solar diffuser and solar diffuser stability monitor, and, at longer wavelengths, a blackbody for thermal band calibration. The instrument will also view dark, deep space as part of its normal scan, and at certain times will image both deep space and the moon as part of the calibration procedure. The inclusion of specific bands for atmospheric sensing on the MODIS instrument provides for dynamic atmospheric correction and allows the estimation of surface directional reflectances in the MODIS land bands (King *et al.*, 1991) (MOD09; 2015).

The use of surface reflectances instead of top-of-the-atmosphere radiances allows a strategy for assembling multidate imagery without relying on specific transforms, such as NDVI, that suppress atmospheric noise. MODIS bands also allow the application of multiple algorithms for the detection of different types of clouds so that cloud-covered pixels may be identified reliably (MOD06). Moreover, in the case of high, thin cirrus clouds, data can be corrected for cloud effects. With atmospherically-corrected and cloud-screened data, measurements can be composited through the fitting of semiempirical BRDF functions (MOD43), allowing calculation of the best-fit nadir reflectance in each of the seven land bands within a 16-day period. In this way, data are obtained that are free of clouds, atmospheric contamination, and angular view and illumination effects.

Rectification has posed particular difficulties for AVHRR processing in producing composited images, analyzing time trajectories of land surface data, and comparing data from multiple time periods to assess change. These problems should be significantly reduced with the in-flight navigation capabilities of MODIS. Geolocation error estimates, discussed in more detail in section 2.4, combined with the growth of pixel size with scan angle, will provide an effective pixel size close to 1-km using the 500-m and 250-m land bands as inputs. Although AVHRR LAC data are commonly composited to 1-km spatial resolution, the data are quite redundant at that scale. For example, normal geolocation error for EDC processing of LAC data is on the order of 2-3 pixels (T. Loveland, personal communication). This error smooths the data greatly and is especially noticeable at abrupt contrast boundaries, such as coastlines. MODIS will thus provide a very significant improvement to 1-km data quality.

3. Algorithm Description

3.1 Overall Algorithm Structure

3.1.1 Practical Description of Algorithm

3.1.1.1 Land Cover Parameter

Figure 1 provides an overview of the Land Cover algorithm logic. In brief, reflected and emitted radiation, as measured remotely through time and over space, are combined with ancillary data to provide a database for distinguishing land covers that includes spectral, directional, spatial, temporal, and collateral information. This database is processed using a neural net classifier.

3.1.1.1.1 Database Assembly and Compositing

In the temporal dimension, remotely sensed data are not retained in full temporal resolution. Rather, the volume of measurements is reduced to a set of summary measurements for each cell in a 1-km grid. In a process similar in principle to the maximum-value compositing of AVHRR data, incoming remotely sensed data are sorted and pruned during a 32-day period to a single composite set of representative values associated with each grid cell. A 32-day period is selected because a number of studies using AVHRR data acquired over large regions have concluded that a period of at least 30 days is required to assemble a dataset that is largely uninfluenced by cloud cover (*e.g.*, Lambin and Strahler, 1994a and 1994b; Moody and Strahler, 1994). The 32-day cycle is keyed to the platform orbit, which repeats in a 16-day cycle and 8-day half cycle.

Remotely sensed inputs to the 32-day database include:

- *Land/Water Flag.*
  Locations permanently covered by water are not investigated by the MODIS Land Cover Classification algorithm. Each record in the 32-day database includes a land/water flag retrieved from an ancillary Level 3 mask. Currently, several such products are under consideration for EOS as a whole.

- *Surface Reflectances.*
  The MODIS BRDF/Albedo product generation executive (MOD43) provides nadir-equivalent reflectances (as an interim product) to MODIS Land Cover in all seven land bands at 1-km resolution. Each observation in this global dataset is the modeled reflectance that would be observed for a given ground location at nadir with the median solar illumination angle for the overpasses over a 16-day period. In effect, this product "composites" all cloud-free observations over the 16-day period using a BRDF model rather than a maximum-value criterion. Thus, two "looks" are available over a 32-day period.

- *Texture Channel.*
  The MODIS BRDF/Albedo Product (MOD43) includes a spatial texture layer. Spatial texture is measured as the ratio of standard deviation to mean of the 250-m surface reflectances falling within each 1-km grid cell during a 16-day period. Studies have demonstrated the utility of this spatial measure in classification of land covers (*e.g.*, Borak and Strahler, 1996). Since texture measures across spectral bands are typically strongly correlated, only the texture for MODIS Band 2 (near-infrared, 250-m nominal resolution) is used. This band provides a strong spectral contrast in regions of varying vegetation cover. As with the reflectance inputs, two looks are available over a 32-day period.

- *Vegetation Index.*
  Maximum-value composited vegetation index is provided by the MODIS Level 3 Vegetation Index Product (MOD34). The spatial resolution of this input is 250-m and its temporal resolution is 16 days. For the 32-day database, the 250-m data from two 16-day products are maximum-value composited, and then spatially averaged to 1-km.

- *Snow/Ice Cover*.
  The MODIS Level 3 Snow/Sea Ice cover Product (MOD33) provides 10-day maps of snow and sea ice cover at 1-km resolution. Only the land portion of this product is used. Because this input product is out of temporal phase with the 32-day database, a linear interpolation across coverage periods is applied.

- *Land Surface Temperature.*
  The MODIS Land Surface Temperature Product (MOD11) provides 8-day composites of land surface temperature. Temperature data are nominally at 1-km resolution. The product is a maximum-value composite.

- *Directional Information.*
  The MODIS BRDF/Albedo Product (MOD43) includes parameter sets describing the fits of semiempirical BRDF models to surface reflectances obtained from MODIS and MISR in the seven land bands. For input to the Land Cover Product, parameters describing the applicable BRDF model type and descriptors of BRDF shape are extracted from this product along with the relevant quality control information. The first parameter, BRDF model type, coarsely preclassifies the type of landscape viewed in terms of the density of intracrown and intercrown gaps. The possible BRDF models differ in terms of their assumptions of leaf density

(low/high leaf area index) and ground object density (sparse/dense). Furthermore, specular components are indicated. The type of land cover viewed is also characterized by the type of BRDF shape provided by the best-fitting BRDF model. This provides for a second parameter, which will be derived in the form of a normalized index characterizing the relative amount of volume to surface scattering in the BRDF observed. If shadows or mutual shadowing of crowns are prominent in a scene, the index will indicate the presence of discrete ground elements (for example shrubs, tree crowns and other rough surfaces). If volume scattering outweighs surface scattering, a rather homogeneous layer of vegetation may be assumed. Relationships between BRDF and land cover are currently under investigation for 3 different ecozones using AVHRR data (see MODIS BRDF/Albedo Product ATBD for details). Generally, the value of using BRDF information in land cover classification derives from the fact that relatively coarse multidirectional class signature boundaries fall across the boundaries of spectral classes. This allows identification of subtleties that may not be apparent in spectral data alone.

The compositing procedure involves examining the overall quality associated with each set of measurements in order to select the best measurements available for the 32-day period. The “best” data are generally those produced with the highest degree of scientific validity according to the MODIS Science Team members responsible for generating the input datasets, *i.e.*, values with optimal quality flags. Often, the degree of validity is related to cloud cover, but it also takes other elements of the production stream into account. Inputs that contain no valid information over a 32-day period are treated as missing data.

Scientifically useful observations from input products that are produced at shorter time steps than 32 days are aggregated to a 32-day time step in the following ways. Land surface temperature is maximum-value composited. Mean weekly snow cover is averaged over the 32-day period. The inputs that are provided by the MODIS BRDF/Albedo Product (nadir reflectances, spatial texture and directional information) are composited in a common manner. For each measurement, the quality associated with each of the two observations available over a 32-day period is examined, and whichever exhibits the higher quality is retained for the 32-day database.

The composited database includes 20 values for each pixel: 7 spectral reflectances and a quality flag; 2 BRDF parameters and a quality flag; single values for spatial texture, surface temperature, vegetation index, and snow cover (each with a quality flag); and a land/water flag. Using a value of 1.5×10^8 sq. km. for the earth's land surface, the total size of the 32-day database is approximately 14 gigabytes.

3.1.1.1.2 Data Reduction

The classification algorithm operates on a sequence of twelve 32-day databases along with an EOS-wide 1-km topographic database to generate seasonal land cover labels. This corresponds to a data volume of approximately 200 gigabytes, which is a prohibitive volume of data to process in the production environment. In addition, some of the input streams are expected to contain redundant information. Therefore, data reduction is applied to the full input measurement space.

Data reduction has been used extensively in remote sensing and classification problems. Probably the most common approach is to employ some sort of linear transformation on the original dataset to produce a smaller set of factors or components (Kauth and Thomas, 1976; Jackson, 1983; Ingebritsen and Lyon, 1985). Most of the original variance is retained with a significant reduction in data volume.

The decision tree classifier (DTC) provides an approach to data reduction that is well known in the pattern recognition literature, but has appeared relatively recently in the remote sensing literature (Michaelsen *et al.*, 1994; Hansen *et al.*, 1996; Borak and Strahler, 1996, section 3.1.4.1.1). The DTC operates in a supervised mode, and thus requires data from training sites. The algorithm employs tree-structured rules that recursively partition the input dataset into increasingly homogeneous subsets based on a splitting rule (Breiman *et al.*, 1984). These subsets are represented as nodes in the tree structure. The top node root) consists of the entire input dataset. Nodes at the bottom of the tree (leaves) are the output cover classes. The hierarchical nature of the classifier thus separates important discriminatory information near the top of the tree from redundant information near the bottom of the tree.

Another method of feature reduction is the use of simple summary variables, such as maximum and minimum values, max-min differences, annually integrated values, etc. These methods have been explored by Lloyd (1990), and more recently, by DeFries *et al.,* 1995.

At present, at-launch data reduction algorithm development is continuing. The pattern recognition field is in rapid development at this time and new algorithms may render old ones obsolete. Final selection of an at-launch data reduction algorithm may be dependent on experience with actual MODIS data. At any rate, the at-launch data reduction algorithm process requires hands-on work at the Scientific Computing Facility, as results will be sent to the DAAC for incorporation into the production environment.

3.1.1.1.3 Classification

In the classifier stage, a sequence of 32-day databases is input to the classifier along with the EOSDIS 1-km topographic database (MOD03). In the post-launch period, the 32-day sequence will involve two years of acquisitions. During the first two years of acquisitions, the sequence will necessarily be shorter. However, we believe at this time that at least one year of observations will be required to make the product. The 32-day databases and ancillary data are then processed by a neural network classifier. Recent applications of neural networks in classification of remotely sensed data were discussed in section 2.2.2. The exact architecture of the neural network classifier will not be determined until the near-launch period. Neural networks and classifiers are evolving so rapidly at this time that it would be unwise to commit to a specific architecture at present. In fact, we may expect the classification algorithm to continue to change during the time span of MODIS data acquisition as techniques improve. A candidate architecture is described in detail in the following section.

For greater processing efficiency and classification accuracy, processing will proceed by continents. Note that the full spectral and temporal resolution of the land cover database may not be needed within each continental region. For example, trials may show that classification accuracy remains unaffected if the annual cycle is represented by a subsample of three or four months. Or, perhaps only a subset of variables within each 32-day composite will be required. In any event, we may predict that the choice of months or variables will vary from continent to continent.

Although a supervised neural network classifier is the candidate classification algorithm at this time, it is possible that an unsupervised classifier or a combination of supervised and unsupervised classification will prove more successful in some ecoregions. A further extension is the first use of a simple classifier (*e.g.*, parallelepiped) followed by a more sophisticated classifier applied only to equivocal pixels. This would increase complexity, but save computation power. These options will be explored as algorithm development proceeds.

3.1.1.2 Land-Cover Change Parameter

The Land-Cover Change parameter will rely primarily on the change vector technique, which will compare pixel-by-pixel the temporal development curve of a set of biophysical and spatial indicators derived from MODIS data. These features are described more fully in section 3.1.1.1.1 above. The change vector technique represents the seasonal dynamic of these indicators by a point in a multidimensional space, with each dimension of this space corresponding to a time-composited observation. Changes in the accumulated value and seasonal dynamic of the indicator between successive years are quantified by a change vector between successive points in the temporal multidimensional space. The magnitude of change is reflected by the length of the vector, while its direction in multitemporal measurement space indicates the timing and nature of the change.

Where a long history of change observations exists, a useful reference standard for change may be a “best-conditions” year, a monthly time trajectory that is constructed for every pixel from an analysis of its historical performance over the period of record. In this process (Lambin and Ehrlich, 1995), the best ecological conditions that occurred throughout the observation period are identified for each pixel by selecting the maximum (or minimum) value of an indicator (*e.g.*, maximum NDVI) for each month in the annual cycle. A best-conditions year is thus constructed that takes into account the ecological and edaphic constraints of every pixel. Any land-cover change can then be expressed with reference to this time trajectory. Note that the reference year will need to be updated at regular intervals to integrate new observations, thus allowing for a better categorization of change processes over multiannual periods. This approach has proven very useful in characterizing land-cover change in Africa using AVHRR GAC data for the 1982-1991 period (Lambin and Ehrlich, 1996c).

Although the change vector technique is the best documented approach thus far to land-cover change at coarse resolution, we also plan to examine other approaches to change characterization. These alternative approaches are identified in section 2.2.4.

3.1.2 Mathematical Description of Algorithm

3.1.2.1 Land Cover Parameter

3.1.2.1.1 Feed-Forward Networks Trained by Backpropagation

While there are a variety of different neural network models, most remote sensing applications have used a supervised, feed-forward structure employing a backpropagation algorithm that adjusts the network weights to produce convergence between the network outputs and the training data. In overview, the neural network classifier is composed of layers of "neurons" that are interconnected through weighted synapses. The first layer consists of the classification input variables and the last layer consists of a binary vector representing the output classes. Intermediate, "hidden" layers provide an internal representation of neural pathways through which input data are processed to arrive at output values or conclusions.

In a supervised approach, the neural network is trained on a dataset for which the output classes are known. In this process, the input variables are fed forward through the network to produce an output vector. During a following backpropagation phase, the synapse weights are adjusted so that the network output vector more closely matches the desired output vector, which is a binary-coded representation of the training class. The network weights, or processing element responses, are adjusted by feeding the summed squared errors from the output layer back through the hidden layers to the input layer. In this fashion, the network cycles through the training set until the synapse weights have been adjusted so that the network output has converged, to an acceptable level, with the desired output. The trained neural network is then given new data, and the internal synapses guide the processing flow through excitement and inhibition of neurons. This results in the assignment of the input data to the output classes. The basic equations relevant to the backpropagation model are presented in Fischer and Gopal (1992).

3.1.2.1.2 Adaptive Resonance Theory Neural Networks

Although the feed-forward back-propagation neural network has been shown to better or at least equal the performance of conventional statistical classifiers in remote sensing applications (see section 2.2.2), this architecture can require lengthy training and can sometimes fail to converge. A newer neural network architecture, relying on adaptive resonance theory (ART), lacks these disadvantages and, in early trials, shows significantly higher accuracies (Gopal *et al.*, 1994). Neural networks employing adaptive resonance theory are designed to be stable enough to preserve significant past learning while still allowing new information to be incorporated in the neural network structure as such information appears in the data input stream. The description of ART networks below follows Carpenter and Grossberg (1987a, 1987b), Carpenter (1989), Carpenter *et al.* (1991a, 1991b, 1992), and Gopal *et al.* (1994).

The ART-1 module of Carpenter and Grossberg demonstrates the essential features of adaptive resonance theory. This module is actually a learning structure that organizes the patterns it receives into a consistent set of responses. In this way, its function is similar to that of an unsupervised classifier. The ART-1 module consists of a two-level network. An input signal in the form of a binary vector (F_0) is received by the first level (F_1) and propagated forward to the second level (F_2) by a set of weights that constitute the long-term memory. A further feature of the F_2 level is that the nodes at the F_2 level interact through lateral inhibition. The result is to produce an F_2 pattern vector in which only the node associated with a single class is significantly activated. This vector is then propagated backward to the F_1 level where it is compared with the original input vector. If the two patterns are close, “resonance” occurs and long-term memory is altered to include the new observation.

If the two patterns differ significantly, the ART module enters a search mode. In this mode, the prior active node at the F_2 level is first disabled. The signal is then propagated forward once again, but since the prior active F_2 node is disabled, a second pattern associated with a different class node is selected. This pattern is then propagated back to the F_1 level and compared with the input vector as before. If the fit is acceptable, then resonance proceeds and the system has “learned” a new input pattern. If not, the second F_2 node is disabled and another attempt to find a good match is made. If cycling in this fashion does not eventually produce a good match, the system adds a new F_2 node and associates the input pattern with it in long-term memory.

Further developments of adaptive resonance theory to neural networks by Carpenter and coworkers include ART-2 and fuzzy ARTMAP. ART-2 modifies ART-1 so that it no longer requires a binary input vector, but instead may accept analog inputs. (A form of scaling is applied to the inputs first, however.) Note that neither ART-1 nor ART-2 are associative memory systems, in which associations between pattern pairs are learned with the ability to be recalled. Fuzzy ARTMAP, however, is such a system.

In fuzzy ARTMAP, two ART-2 modules (ART_A and ART_B) are connected together through an associative learning network called a map field (ART_{AB}). In the training phase, input vectors and desired output vectors are presented as pairs to F_2 and ART_B respectively, and the outputs of the ART modules are associated by the map field ART_{AB}. If a mismatch occurs, F_2 is placed in search mode to find, and possibly learn, a better choice. Or, a new node may be added to the F_2 layer in F_2 that is a better predictor of the desired output. As noted previously in section 2.2.2, the fuzzy ARTMAP classifier has performed very well in several trial applications to monthly-composited AVHRR LAC data as well as to Thematic Mapper data.

Fuzzy ARTMAP is an example of newer approaches to the use of neural networks in pattern classification. Undoubtedly neural network architectures will continue to develop in the pre- and postlaunch periods. We plan to monitor developments in neural network architectures, and as candidate architectures are produced, we will test them for application to the MODIS Land Cover Parameter.

Note that both of the above techniques operate in supervised classification mode. Although the supervised approach is well-suited to a situation where training data are aggregated into test sites, it is possible that unsupervised approaches may be more desirable for some regions or situations. In that event, a number of unsupervised neural network pattern classifiers are available and have been used on remotely sensed data with good results (*e.g.*, Hara et. al, 1994).

3.1.2.3 Decision Tree Classifiers

Decision trees classifiers recursively partition data into related or homogeneous subregions based on a set of decision rules. The structure of the tree consists of a root node, intermediates nodes or splits, and terminal nodes (leaves). Input data at the root node is subdivided at decision nodes based on univariate and/or multivariate decision rules (Brodley and Utgoff, 1992, 1995). Although decision trees have only recently been applied to remote sensing data, they offer tremendous potential for classification and feature selection. Decision trees can either be applied in independently or coupled with other analytical procedures in hybrid classification models.

Research into the applicability of decision trees to MODIS Land Cover continues to be conducted at both site and global scales. Lloyd (1990) used a binary decision tree classifier using multitemporal phenological indexes or metrics derived from a time series of NDVI data to stratify vegetation phenology classes. A time series dataset was used to examine feature selection and land cover classification for a MODIS-like scenario in the semiarid environment of the Walnut Gulch/Cochise County site (Borak and Strahler, 1996, and section 3.1.4.1.1). The dataset consists of numerous input fields derived from an intraannual sequence of seven Landsat TM acquisitions, along with ancillary elevation information. A decision tree classifier selected the features that provided the best discrimination among land cover types. Three classification algorithms were applied to the reduced feature space: the decision tree itself, a maximum-likelihood classifier and an artificial neural network (Fuzzy ARTMAP). Results indicated that decision tree classifiers are useful tools for extracting essential features in data sets of high dimensionality, and that the neural network performed best on the reduced set of features.

Brodley *et al.* (1996) applied univariate, multivariate and hybrid decision trees to the global 1 degree AVHRR NDVI dataset, the Conterminous US 1-km AVHRR NDVI dataset, and a TM data set for Lake Tahoe, California, and compared them to a linear classifier. All decision trees outperformed the linear classifiers on all datasets. Hybrid trees, where different classification algorithms are used in different subtrees of a larger tree, were superior due to their ability to better resolve complex relationships among feature attributes and class category labels. The empirical results of these studies indicate that research should continue in the application of decision trees to land cover classification.

3.1.2.2 Land-Cover Change Parameter

The change vector analysis method of identifying land-cover change has been presented by Lambin and Strahler (1994b). This method assesses change by calculating the distance between the location of an indicator variable in multitemporal space at two different time periods on a per-pixel basis. For example, if the indicator were monthly composited NDVI values, and it was desired to assess change between two subsequent years, the difference between the location of the NDVI vector in 12-dimensional (monthly) space for the two years would be calculated. The magnitude of this difference, or change vector, is representative of the magnitude of the change, and the direction of the change vector relates to the type of change. The type of change is then characterized by the segment of the multidimensional temporal space into which the vector has moved. This approach (1) allows quantification of the intensity of change;(2) allows classification of the type of change; (3) is based on a historical database for each pixel; and(4) is mathematically simple.

The change vector analysis procedure is simply represented by Lambin and Strahler (1994b). In the case where a single remotely sensed variable is measured (for example the NDVI) the temporal state of the land cover can be represented by the location of that variable in a multidimensional space, where each dimension represents one of the time periods for which the variable was measured. For example, the location of the variable can be represented by the vector:

$$\mathbf{p}'(i,y) = \left[I(t_1) \quad I(t_2) \quad \ldots \quad I(t_n) \right]$$

where $\mathbf{p}(i,y)$ is the multitemporal vector for pixel i in processing period y, I are the values of the variable of interest for pixel i at time periods t_1 to t_n where n is the number of time periods at which the variable was measured. The vector magnitude represents the integral of the variable over time periods 1, 2, . . ., n and the direction of the vector represents the seasonal pattern. Any change in the state of a pixel's land cover between processing periods y and z is defined by a change vector:

$$\mathbf{c}(i) = \mathbf{p}(i,y) - \mathbf{p}(i,y)$$

The magnitude of the change vector is simply taken as the Euclidean distance d between the two vectors:

$$d^2 = \mathbf{c}'(i)\mathbf{c}(i)$$

In work by Lambin as cited above, it has been shown that the magnitude and direction of the change vector are related to the intensity and type of change process, respectively.

A further extension is to explore the Mahalanobis distance, $d_{M:}$:

$$d_M^2 = \mathbf{c}'(i)\mathbf{V}^{-1}(i)\mathbf{c}(i)$$

where $\mathbf{V}(i)$ is a variance-covariance matrix of change vectors that quantifies the distribution of change vectors as observed over time within a region. This distance scales the magnitude of a vector by the variance normally observed in its direction of change.

3.1.3 Test Sites and Training Data

The MODIS Land Cover and Land-Cover Change Parameters require ground information for training and validation. This information will be obtained for a network of global test sites described in the following sections. We are currently completing the development and population of a Global Validation and Test Site database to manage this site information.

3.1.3.1 Site Hierarchy

- *Regional Sites.*
  We project that five Regional Sites will provide an appropriate scale for algorithm testing and development, as well as at-launch and post-launch validation. Two currently-identified Regional Sites are Central America and the Caribbean (Central America Terrestrial Ecology Study), and the Western U.S. and Canada. Data for these sites include multitemporal 1-km AVHRR, single-date TM data for as much of the region as possible, ancillary environmental and climatological data, and field inventory.

- *Instrumented/Tower Sites.*
  As part of the larger MODLand, MODIS, and EOS activities, a network of fully-instrumented tower sites is being developed. We will use these sites for pre-launch evaluation as they become available and for continuous observation in the post-launch era. They will be visited during field campaigns in order to map and parameterize local land cover types and observe change.

- *Intensive Study/Continuous Observation Sites* .
  A critical component to training and validation is a stratum of Intensive Study/Continuous Observation Sites. These sites are critical to testing algorithms, process and products. In the pre-launch period, they represent areas for which temporal sequences of TM imagery are available. TM provides spectral coverage similar to six of the seven MODIS land bands and can be convolution-resampled to provide data at 250- and 500-m resolution (Barker *et al.*, 1992). In this way, the multispectral and multitemporal information content of MODIS imagery can be simulated before the actual acquisition of MODIS data. Also useful is AVHRR imagery. Although the spectral bands of the AVHRR correspond to only two MODIS land bands, the instrument provides a source of thermal data registered to the VNIR radiances and further simulates the angular geometry of MODIS and its temporal repeat pattern for studies of cloud cover and compositing.

  The Walnut Gulch/Cochise County site (discussed in section 3.1.2.1.3) is an example of a site of this type. Research at Walnut Gulch has included feature selection and land cover classification of a MODIS-simulated dataset derived from multitemporal, spatially-degraded Landsat TM data (Borak and Strahler, 1996 and in preparation and Section 3.1.4.1.1). These studies have clearly demonstrated the applicability of both neural network and decision tree classifiers to a specific land cover classification problem. In addition, the study has helped us identify processing issues that the Science Computing Facility (SCF) will face in the MODIS processing era. We are continuing analysis and field inventory at this site in cooperation with the SALSA Field Campaign.

  Additional Intensive Study/Continuous Observation Sites we have identified include BOREAS, Olancho, Honduras; the Yucatan; Glacier National Park; Plumas National Forest; Harvard Forest LTER; Hubbard Brook LTER; Virginia Coast LTER; and Hapex-Sahel.

- *Field Validation Sites.*
  We are developing 100-150 sites at which remote sensing-derived data are validated through field inventory and observation. These Field Validation Sites are a subset of our global Remote Sensing Sites. Each site consists of one or more large polygons of homogeneous land cover that can be used in training and validation of any land cover product. Attached to the polygon are both image segments from selected remotely sensed data sources (AVHRR, TM, MODIS, ASTER, ETM, Lewis, etc.) and field measurements of ecological and biophysical parameters. At first, field data will be acquired from existing sources, such as field sampling in Rapid Ecological Assessment studies carried out by the conservation community. In the near term, field data will be acquired specifically for field validation sites as travel opportunities present themselves. Friends and collaborators will also be tapped to acquire data. Over the lifetime of the MODIS instrument program, these sites will constitute a very useful set of consistent data for parameterization of many ecosystem characteristics and validation of alternative classifications.

- *Remote Sensing Sites.*
  Global training, testing and validation require that a network of remote sensing observation sites be developed. As in the Field Validation Sites above, the database will consist of polygons or grid cells of homogeneous land covers. However, parameters (*e.g.*, IGBP land cover class) will be inferred from manual interpretation of high-resolution imagery. This activity is currently underway using, Landsat TM and MSS and SPOT data, at both Boston University and the University of Maryland, College Park. The goal is a representative network of approximately 500 sites.

3.1.3.2 Site Selection Criteria

Since the test site database will need to be representative of the earth's diversity of land covers and types of land cover change, a site selection strategy is required. Five criteria will be exercised.

- *Global Representativeness.*
  The nature of the earth's land cover at a location is dependent on the independent variables of climate, physiography, human activity, and ecological potential. As these factors vary, so do the land cover types, their modes of change, and the ecological and biophysical parameters that characterize them. Thus, any land cover sample that is to be globally representative must cut across these factors to sample their variation in an independent fashion. For example, it would be possible to select sites stratified by latitude, continentality, solar irradiation, precipitation, elevation, or slope. This stratification would provide the best independent means of determining and sampling the overall “representativeness” of the earth from a climate/vegetation perspective. Yet these stratification factors are themselves correlated into broad regions with individual patterns of land covers and change processes.

  Regionalization is the approach to stratification whose objective is to increase classification accuracy by subdividing the globe into regions that minimize within-stratum variance and maximize between-stratum variance. Stratification by regions is powerful and effective, but requires definition of regions. This, however, is dependent on how MODIS data organize themselves as they are forced by the independent factors above and must thus regionalization must await the post-launch era of MODIS earth observation.

- *Change Detection Sites.*
  Land-Cover Change Parameter training and validation requires that the site network be representative of global, regional and local change processes due to both natural and anthropogenic factors. Change criteria include phenological class (seasonal grassland, deciduous forest), anthropogenic (urbanization, agriculture, conversion, biomass burning), interface (land/snow, land/water), biotic (insect and pathogen), and hydrologic (seasonal inundation) representation.

- *Indicator Sites.*
  Early warning and indicator sites are necessary for monitoring processes that may be indicative of climatological or other changes. These sites include, for example, high-elevation spruce-fir forests and ecological ecotones.

- *Critical Sites/ Hotspots.*
  A number of sites are to be included in the network because of their particular conservation, political, economic and/or social significance. Rondonia is an example.

- *Statistical Criteria.*
  A proper global sample would be drawn on a statistical basis using probabilities of sample selection associated with each potential site. Although it may not be possible to draw such a sample given practical limitations, statistical criteria that could be exercised informally include sample stratification by geographical or regional type; stratification by land cover or change process after these have been determined in the postlaunch period; sample selection for variance reduction; and sample selection optimized for change detection. Issues in global sample selection for land cover classification accuracy assessment are described in more detail in Belward (1996).

3.1.3.3 Site Data Sources and Institutional Cooperation

A global site network is an ambitious endeavor that requires cost sharing and inter-institutional cooperation. Test sites are promoted to take advantage of cost savings by using data that have already been generated and where research is continuing. This site network is being coordinated within MODLand, EOS and the larger remote sensing community, especially IGBP. Boston University is working with other MODLand teams in developing site data, including work with the Snow/Ice Team in Glacier National Park and a winter 1996 field and MAS data acquisition campaign for Hubbard Brook LTER, New Hampshire.

One important source of test site data is the Landsat Pathfinder Global Land Cover Test Sites (GLCTS) program, for which data are now being assembled at the EROS Data Center. GLCTS was proposed by an informal network of researchers largely connected to

MODIS, EOS and/or the IGBP. Each site is known to have some associated ground information on local land cover or related data. For each GLCTS site, a database of remotely-sensed imagery that includes Landsat MSS, TM, and AVHRR data is being assembled by the GLCTS program. Landsat images include both recent and historical acquisitions (for change detection). AVHRR images include LAC data in a 500-km by 500-km window centered on each test site; they are being acquired as part of the data acquisition phase for the AVHRR global 1-km dataset, and acquisitions are expected to continue into the future along with this program. Databases have been completed at eight sites. The present list of GLCTS test sites includes 130 locations globally, although there is a commitment now for only database development at some 30 sites. Although the completed sites are data-rich, there is no consistency in the level of land cover information that has been derived at each site. It will therefore be necessary to extract land cover data for the GLCTS sites to make them usable for MODIS Land Cover training and validation.

Another source of test sites is the IGBP-DIS Global 1-km Land Cover Database project, which will identify a global set of confidence sites to provide local-scale land-cover information for both methodological development and accuracy assessment of the 1-km land cover database (IGBP-DIS, 1994). The project anticipates 40 to 100 such “core” sites, many of which will likely coincide with those of the Landsat Pathfinder. IGBP has also proposed to validate the IGBP AVHRR global land cover classification using statistical sampling based on over 400 remote sensing observation sites. These data will be available and may be easily applied to validation of the MODIS product, even though they are not designed specifically for that purpose.

Other sources of test sites are international field experiments, such as the Boreal Ecosystem-Atmosphere Study (BOREAS) in boreal Canada; the SALT transect (Savanna on the Long-Term) in West Africa; or the French test sites for the POLDER (Polarization and Directional Earth Reflectance) instrument.

3.1.3.4 Site Databases

Test site data need to be compiled, analyzed and stored in an efficient information management system. We are currently developing several databases that support our global test site network.

- *Global Thematic Database.*
  We are assembling a database of ancillary global data to support analysis of test site location and as collateral data for classification and validation. These data include vegetation maps and databases of Matthews, Olson, Kuchler, and Loveland; climate maps and databases of Köppen and Strahler; and Land-Surface/Land-Form maps and databases (TBD).
- *Test Site Databases.*
  We are compiling field and ancillary thematic and remotely sensed data for our Regional sites, Instrumented/Tower Sites, Intensive Study/Continuous Observation Sites, Field Validation Sites, and Remote Sensing Sites.

· *Validation and Test Sites (VATS) Parameter Database.*
To make the site data useful, variables and parameters need to be extracted and compiled in a site database; this is the operational training and validation data. This database is the repository for remote sensing and other variables extracted and tracked for individual plots and sites. It includes field data coupled to specific points and polygons.

3.1.4 Prelaunch Algorithm Development and Validation

3.1.4.1 Land Cover Parameter

The primary objectives of algorithm development and validation in the prelaunch phase are to: (1) validate the general and specific neural network architecture to be used in the classifier; (2) select the best sources of information for classification from spectral, temporal, spatial and directional domains; and (3) identify and refine the compositing criteria used in assembling the database. These objectives require test site data that simulate the information content of MODIS imagery as closely as possible. For this purpose, the best source of such data is a registered sequence of TM images acquired during an annual cycle.

TM is preferred to other high-spatial resolution sensors because TM spectral bands are similar to those of the MODIS land bands. To simulate MODIS imagery, the MODIS Science Data Support Team has written a program that executes a spatial convolution procedure on TM imagery to provide outputs at 250-m, 500-m and 1-km resolutions (Barker *et al.,* 1992). Assembling a database of such imagery over an annual cycle adds the multitemporal dimension. Texture information can also be generated from such a sequence, providing the spatial dimension. As directional information is lacking, multitemporal TM data can provide three of the four MODIS information domains: spectral, temporal, and spatial.

3.1.4.1.1 Walnut Gulch/Cochise County Classification Study

As a part of our prelaunch validation work, a MODIS-like dataset was created and used to examine feature selection using decision trees and land cover classification using both decision trees and neural networks for a portion of Cochise County, Arizona, including the Walnut Gulch study area (Borak and Strahler, 1996). The dataset consisted of 57 input fields per pixel derived from an intraannual sequence of seven Landsat TM acquisitions merged with ancillary elevation information from a 30 arc-second EDC digital elevation model. The TM database was subsetted so that a region of approximately 13,000 km^2 could be examined in detail.

Land cover reference data for training and testing were derived from a statewide base map of land covers prepared for the Gap Analysis Project (GAP). These data were converted from vector polygons to a 30-m raster by the University of Arizona and generously made available to us. Within the study region, 42 class labels were present. These classes were relabeled into the IGBP scheme; 12 of the 17 IGBP classes were represented. No accuracy data were available for the database.

Data used were resampled to a common 960-m grid that approximated the spatial resolution of the MODIS Land Cover Product. The satellite spectral data were convolved with a TM-to-MODIS Gaussian blur filter using a fast Fourier transform (Barker *et al.*, 1994) in order to simulate MODIS imaging characteristics. In addition, a 240-m NIR channel was created for the derivation of a texture index. The aggregated spectral data were corrected to surface reflectance employing the MODTRAN atmospheric characterization algorithm (Berk *et al.*, 1989), which was calibrated by ancillary meteorological data. After atmospheric correction, NDVI and SAVI were produced at 960-m resolution and the MODIS texture measure, taken as the standard deviation of all 240-m NIR values within the 960-m pixel, was calculated. Finally, the elevation dataset was registered to the resampled satellite data.

The reference dataset was registered to the full-resolution satellite data using a concurrence of land cover types with obvious features present in the images. To produce pixels at 960-m resolution, the coarse resolution land cover was determined by applying a plurality voting rule to the fine-resolution labels.

In the data analysis phase, a decision tree classifier was first applied to select the inputs, or features, from the full measurement space that were the more significant discriminators of land cover type. Three classification algorithms were then applied to the reduced feature space: the decision tree, a maximum-likelihood classifier (MLC) and the fuzzy ARTMAP artificial neural network. To examine the effect of sampling of training data on accuracy, two sampling approaches were employed:training data were sampled proportionally to class size in the reference map, and training data were sampled with fixed frequency for all cover types. In each case, 20 percent of observations in the database were used for training, while all observations were used for accuracy assessment.

In the proportional case, the decision tree identified 14 significant features. These included each of the six spectral bands, SAVI, and NIR texture on at least one date, as well as elevation. For the fixed-size case, 10 features were selected that included five spectral bands, SAVI, image texture, and elevation. Thus, all inputs contained useful information at some point in the time series. With respect to data compression, the decision tree reduced the input feature space by 75 percent and 83 percent for the proportional and fixed-size sampling approaches, respectively. As a control, the full feature space was retained for one neural net classification trial.

In the classification phase, five classifications were carried out: maximum likelihood with and without prior probabilities; decision tree; and fuzzy ARTMAP on both the full and reduced feature set. Table 3 presents the results expressed as classification accuracies. The neural net classifier clearly had a significant performance advantage in comparison to the tested alternatives. For the proportional case, fuzzy ARTMAP produced an overall accuracy that was 4.9 percent above its closest rival, and for the fixed-size approach, this difference increased to 9.2 percent.

Note also that the neural net classifier performed about the same using both the full and reduced feature sets. This indicates that feature reduction did not degrade accuracy. It also demonstrates the ability of the fuzzy ARTMAP classifier to handle data redundancy without difficulty.

Table 3. Classifier Comparison, Walnut Gulch/Cochise County MODIS Simulation Study

	Sampling Proportional to Size		*Fixed-size Sampling*	
Classifier	***Accuracy, percent****	***Kappa statistic***	***Accuracy, percent****	***Kappa statistic***
Maximum likelihood	56.3	0.401	54.4	0.306
Maximum likelihood, with prior probabilities	71.1	0.518		
Decision tree	73.5	0.446	47.2	0.306
Fuzzy ARTMAP	78.4	0.609	63.6	0.450
Fuzzy ARTMAP, all features	77.3	0.583	64.1	0.455

**Includes mixed pixels and errors in GAP database.*

For all classifiers, the proportional sampling strategy produced higher accuracies, a result that demonstrates the importance of information about the relative size of classes on overall accuracy. However, the fixed-sample strategy produced a somewhat more even distribution of errors of omission and commission for individual classes.

The table also shows that overall accuracies for this study are about 10 percent lower than those normally achieved in this type of land cover classification, for example with the 1-km AVHRR NDVI dataset described in the following section. We attribute this to inaccuracies in the “ground truth” database. In the preparation of any map of thematic polygons, generalization is required to produce smooth, coherent, and consistent boundaries. Small patches or clusters of inliers and outliers are removed. However, when classifiers are applied to observations on a raster grid, the results typically, and correctly, include these patches. After relabeling in the IGBP scheme, the GAP map was quite generalized, showing only large polygons, even at 960-m resolution. For the entire study area, exceeding 12,000 km^2, only about 200 polygons were present, providing a mean size of about 650 MODIS-sized pixels per polygon. This statistic aptly demonstrates the simplified nature of the ground truth data.

3.1.4.1.2 Neural Net Classification of Global 1° AVHRR Dataset

Although MODIS-like datasets are the most helpful for developing algorithms in the prelaunch period, it is also useful to test candidate classifiers on existing global datasets. Accordingly, we repeated the procedure of Townshend and DeFries (1994) in classification of a coarse resolution (1° in latitude and longitude) multitemporal composited AVHRR NDVI dataset (Los *et al.*, 1994), substituting a fuzzy ARTMAP neural net classifier (see section 3.1.2.1.2) for their maximum likelihood classifier.

Townshend and DeFries used eleven broad vegetation cover types rather similar to those of the IGBP (Table 1). The training sites for each of the cover types were identified based on agreement between three existing datasets of land cover (Matthews, 1993; Wilson and Henderson-Sellers, 1985; Olson, *et. al.*, 1983). That is, all 1° cells for which the class assignment by each of these three datasets agreed were taken as pixels with known class labels. These pixels were then used for training and testing. It is interesting to note that the three data sets agree on cover types for only 26 percent of the total land surface. The dataset consists of a total of 3398 pixels.

For this work, Townshend and DeFries tested both a calendar sequencing of NDVI values and a resequencing beginning with the peak value in order to standardize seasonal effects. That is, in resequencing, the first NDVI value is the maximum (regardless of the month in which it occurs), the second value is the value that occurs in the month following the month of the maximum, and so forth. They also explored several ways of incorporating latitude into the classification process, for example by stratifying the land surface into latitudinal strips or subdividing the training areas by latitudes. Using a maximum likelihood classifier (MLC) on the entire set of training data, they concluded that both incorporating latitude as a variable in the classification and resequencing the NDVI values provided more accurate results.

Using the same datasets (generously made available by Townshend and DeFries) we reclassified the labeled pixels using a fuzzy ARTMAP network. The network was trained using 80 percent of these pixels and testing was done on the remaining 20 percent. Paralleling the work of Townshend and DeFries, we also investigated the effect of latitude on classification accuracy. In one case, latitude was scaled from 0 at the north pole to 180 at the south pole, whereas in the other latitude was omitted entirely.

Table 4 presents pixel counts for training and testing data as well as accuracies, expressed as percents, for the fuzzy ARTMAP classifier exercised on data including and omitting latitude. The results are averages from five different neural networks built from different random presentations of the training data. Overall accuracy is about 10 percent higher when latitude is used directly in the classifier, confirming the conclusions of Townshend and DeFries. Note also that the neural net classifier outperforms maximum likelihood by about 8 percent. The table also shows the results of classifications using decision trees, which are discussed in the following section.

It is also interesting to observe the variation in global classification results between classifiers. Table 5 shows pixel counts and percentages for global classifications using two of the classifiers, fuzzy ARTMAP and the hybrid decision tree (discussed further below). In general, the global percentages are similar to those of the input training sites, emphasizing the importance of class size within the training set on the classification outcome (see section 3.1.4.1.1 for more discussion of this point). This behavior is especially apparent for broadleaf deciduous forest/woodland, which in neither of the global classifications exceeds 1 percent of the total pixels. In this case, the very small number of training pixels constrains the classifiers to limit the number of classified pixels as well. The results point up the importance of obtaining training data that is representative of global land cover in all respects (see section 3.1.3).

Table 4. Classification Results, AVHRR Composited NDVI 1° Dataset

Land Cover Category	*Total Training Pixels Available, Townshend and DeFries*	*Neural Net ARTMAP Classifier*				*Maximum Likelihood and Decision Tree Classifiers*						
		Pixel Count		*Classification Accuracy, percent*		*Pixel Count*			*Classification Accuracy, percent*			
		Training (80 %)	*Testing (20%)*	*With Latitude*	*Without Latitude*	*Training (70%)*	*Pruning (20%)*	*Testing (10%)*	*Maximum Likelihood*	*Univariate Decision Tree*	*Multivariate Decision Tree*	*Hybrid Decision Tree*
Broadleaf evergreen forest	628	502	126	97.6	91.3	440	125	63	84.0	95.7	96.1	96.3
Coniferous evergreen forest/-woodland	320	256	64	68.8	53.1	224	64	32	69.0	81.0	80.0	78.1
High latitude deciduous forest/woodland (mostly larch)	112	90	22	86.4	90.9	78	23	11	100.0	85.4	92.7	99.1
Tundra	735	588	147	95.9	94.6	514	147	74	86.0	94.5	94.7	95.1
Mixed deciduous and evergreen forest/woodland	57	46	11	63.6	72.7	40	11	6	40.0	34.0	56.0	58.0
Wooded grassland	212	170	42	88.1	45.2	148	43	21	95.0	82.9	86.7	86.7
Grassland	348	278	70	80.0	68.6	243	70	35	35.0	70.6	66.2	70.3
Bare ground	291	233	58	100.0	89.7	204	58	29	100.0	95.2	95.6	97.3
Cultivated	527	422	105	74.3	79.5	369	105	53	81.0	76.6	78.5	80.2
Broadleaf deciduous forest/-woodland	15	12	3	0.0	0.0	10	3	2	100.0	0.0	0.0	0.0
Shrubs and bare ground	153	122	31	80.6	77.4	107	31	15	100.0	84.7	86.0	92.7
Pixel total	*3398*	*2719*	*679*			*2377*	*680*	*341*				
Overall accuracy				*86.6*	*76.7*				*78.8*	*85.6*	*86.4*	*87.7*

Table 5. Global Classification Pixel Counts and Percentages						
			Global Classification			
	Training/Test Data		*Fuzzy ARTMAP (with latitude)*		*Hybrid Decision Tree*	
Land Cover Category	*Pixel Count*	*Per-cent*	*Pixel Count*	*Per-cent*	*Pixel Count*	*Per-cent*
Broadleaf evergreen forest	628	18.5	1770	10.8	2130	13.0
Coniferous evergreen forest/woodland	320	9.4	1805	11.0	2145	13.1
High latitude deciduous forest/woodland (mostly larch)	112	3.3	640	3.9	921	5.6
Tundra	735	21.7	2723	16.6	2415	14.7
Mixed deciduous and evergreen forest/woodland	57	1.7	541	3.3	250	1.5
Wooded grassland	212	6.2	1675	10.2	1264	7.7
Grassland	348	10.2	2287	13.9	2527	15.4
Bare ground	291	8.6	1502	9.1	1510	9.2
Cultivated	527	15.5	2857	17.4	2705	16.5
Broadleaf deciduous forest/woodland	15	0.4	26	0.2	19	0.1
Shrubs and bare ground	153	4.5	592	3.6	533	3.2
No class label			1	0.0	0	0.0
Total	*3398*	*100.0*	*16419*	*100.0*	*16419*	*100.0*

3.1.4.1.3 Decision Trees

As noted in section 3.1.2.1.3, decision tree classifiers present an alternative to neural networks that we are also exploring for use in preparing the land cover product. Friedl and Brodley (1996) recently tested the performance of decision tree algorithms for land cover classification. Three different decision tree algorithms were tested on several different training data sets. The decision tree algorithms included univariate decision trees (UDT; Breiman *et al.*, 1984), multivariate decision trees (MDT; Brodley and Utgoff, 1995) and hybrid decision trees (HDT; Brodley, 1995). Univariate decision trees test a single feature at each internal node, whereas multivariate decision trees use linear combination tests to define the splitting criteria at each internal node. Hybrid decision trees employ multiple classification algorithms within the framework of a single decision tree structure.

The performance of these algorithms was evaluated using three different datasets with different spatial, spectral and temporal properties. The first was the 1° AVHRR composited NDVI dataset of Los *et al.* (1994) that was used by Townshend and DeFries (1994) and Gopal, Woodcock and Strahler (1996) as described above. Training labels for these data were derived by Townshend and DeFries (1994). The second dataset tested was derived from the 1990 Conterminous US AVHRR Dataset compiled at EROS Data Center (Eidenshink e*t al.*, 1992) and consists of a time series of maximum NDVI values during each month of the growing season in 1990. Class labels were assigned to these data by reclassifying the labels provided by Loveland *et al.* (1991) to the IGBP classification (see section 2.3). The data used here were extracted at 10,000 random locations, and exclude water bodies. The final dataset was composed of a random sample of roughly 2000 values of raw Landsat Thematic Mapper (TM) data acquired over a forested area surrounding Lake Tahoe, California. These data were extracted from a single image and include all TM bands except band 6 (thermal). Class labels were assigned using a combination of automated classification procedures that incorporated the use of ancillary data, manual labeling using field data, and aerial photography (Woodcock *et al.*, 1994a, b).

To evaluate the classification performance of each of the decision tree algorithms identified above, a set of ten cross-validation runs were performed using each classification algorithm to classify each of the datasets. To do this, each dataset was split into three parts: 70 percent training, 20 percent pruning, and 10 percent testing. In this way, the trees were estimated, pruned, and evaluated using independent data for each step. This procedure was repeated to generate ten versions of the data with different random combinations of training, pruning, and testing data. To provide a baseline of the performance of the decision tree algorithms a parallel cross-validation procedure was performed using both a maximum likelihood and linear discriminant function classifier. The classification accuracies were then reported as averages across the ten cross-validation runs and are presented below in Table 6. These results show that the decision tree algorithms consistently and significantly outperformed more conventional classification algorithms.

Table 6. Classification Results, Decision Tree Classifiers

Classification Method	***NDVI 1°***	***NDVI 1-km***	***TM***
Univariate decision tree	85.6	71.0	75.2
Multivariate decision tree	86.4	71.7	75.9
Hybrid decision tree	87.7	80.1	76.0
Linear discriminant functions	78.7	51.7	70.6
Maximum likelihood classification	78.8	62.2	69.0

Table 4, presented earlier, provides per-class accuracies for the NDVI 1° dataset. Although for some classes maximum likelihood gave more accurate results, superior performance on the larger classes gave decision trees an overall accuracy advantage of 7-9 percent.

It is notable that classification accuracies are in general lower for the NDVI 1-km and TM datasets than for the NDVI 1° dataset. At 1-km resolution, the data are significantly noisier than at 1°. The NDVI 1-km dataset suffers from multidate registration inaccuracies and is also not fully corrected for atmospheric effects. In contrast, the 1° dataset is filtered and smoothed (Los *et al.*, 1994) as well as averaged over a much larger grid cell size. Note that both multidate registration and atmospheric correction will be much better for MODIS than for the AVHRR NDVI 1-km dataset. Noise and mixed pixels are also characteristic of the TM dataset (Woodcock *et al.*, 1994a, b).

In addition to this assessment of decision tree performance, Brodley and Friedl (1996) developed a consensus filter technique based on machine learning theory to filter training data for mislabeled observations. Given the importance of training data quality to the performance of both neural net and decision tree classification algorithms, this method is of potentially high utility for the implementation of operational land cover and land-cover change algorithms. Specifically, the technique provides an automated and objective method for tagging mislabeled observations in training data that are input to supervised classification algorithms.

The consensus filter technique uses an *n*-fold cross validation procedure using a set of *m* different classification algorithms. To this end, the algorithm first splits the training dataset into *n*-subsets. For each of the *n*-subsets, each of the *m* different classification algorithms are then estimated using the *n*-1 other subsets. The *m* classification algorithms are then applied to the unseen data. Each classifier identifies a sample as potentially mislabeled if it does not classify the sample correctly. An observation is tagged as being incorrectly labeled by the filtering procedure if all of the *m* classifiers identify it as being mislabeled (*i.e.*, a "consensus"). At the end of the *n*-fold cross-validation, each sample in the training data has been tagged as being either correctly or incorrectly labeled.

The basic notion behind this procedure is that if several different classification algorithms all classify a training sample differently from its assigned label, there is a high probability that the observation is mislabeled. The theoretical basis for this technique is founded in the observation that different classification algorithms have different biases (Brodley, 1995). In other words, different classification algorithms may classify individual observations differently because the observations appear to be outliers relative the specific algorithm being used. By applying multiple classification algorithms with different biases and requiring a consensus among the different algorithms, the probability is increased that an observation that is tagged as being mislabeled actually is mislabeled. In this context, because the procedure requires a consensus among the *m* classifiers, it is designed to be conservative.

To test this procedure, we simulated labeling errors in training data by randomly introducing error between classes that are likely to be confused in real data (*e.g.*, grassland versus wooded grassland). To do this we used the training dataset generated by DeFries and Townshend (1994) for 1° the AVHRR composited NDVI dataset (Los *et al*., 1994). Results from this analysis show that the procedure is capable of detecting and removing

Table 7. Effect of Noise and Consensus Filtering on Classification Accuracy (percent).

	Classifier					
	K-Nearest Neighbor		***Linear Discriminant Functions***		***Decision Tree***	
Induced noise level	***Unfiltered***	***Consensus-filtered***	***Unfiltered***	***Consensus-filtered***	***Unfiltered***	***Consensus-filtered***
0	87.3	N/A	78.6	N/A	85.6	N/A
5	84.4	87.8	76.7	78.9	84.4	86.1
10	81.8	86.4	77.5	79.2	80.1	85.1
20	75.8	83.1	70.2	78.1	75.2	81.8
30	68.6	75.2	63.4	74.0	67.4	74.2
40	58.4	59.9	49.0	54.2	56.9	59.6

fairly substantial levels of noise in training data (Table 7). Note also that the procedure is amenable to use in training the classifier for the MODIS land cover product, since it operates on training data rather than the entire global dataset.

3.1.4.2 Land-Cover Change

Prelaunch algorithm development of the Land-Cover Change Parameter has three primary objectives: (1) to validate the multitemporal change vector technique at broad spatial scales (continental) and over a decade or more of observations; (2) to refine the logic for land-cover change characterization; and (3) to define the linkage between the land-cover change technique and ancillary data (thematic information and high resolution information) for a more detailed monitoring of "hot spots"—areas of rapid change. This work will extend the research of Lambin and Strahler (1994a, 1994b) on the application of change vector analysis to AVHRR LAC data from west Africa.

The objectives above require a long time series of high temporal-frequency satellite observations. Only AVHRR GAC data meet these requirements. There are two important sources of GAC data for this purpose. First is the AVHRR Pathfinder dataset, comprised of 12 years of daily cloud-screened GAC data that are calibrated, corrected for ozone absorption and Rayleigh scattering, and registered to a map projection. Another important time sequence of AVHRR data is the GAC dataset of the African continent produced by the Monitoring Tropical Vegetation Unit of the European Union's Joint Research Center in Ispra, Italy. This dataset has already been used extensively in developing the change-vector technique by Associate Team Member Eric Lambin (Lambin, 1995; Lambin and Ehrlich, 1995, 1996a, 1996b).

It is important to note that because the spectral bands of the AVHRR are limited, these data do not provide the full information content of the MODIS database. However, AVHRR data do provide the temporal signal that is expected to be most important in the detection and characterization of land-cover change. In the prelaunch period, algorithm development activities are focusing on understanding and exploiting the information content of the AVHRR temporal signal.

Specific near-term activities in algorithm development of the Land-Cover Change Parameter include: (1) continuing analysis of change vectors over ten years of African NDVI and surface temperature data; (2) validating the change magnitude and change processes that are detected using a variety of sources, such as FAO reports, FEWS Bulletins, NCAR Climate Impact Maps, published reports on land-cover change, and high resolution data analysis; and (3) defining typical temporal patterns of land-cover change to establish the basis of a future classification system for land-cover change processes.

3.1.5 Sources of Error and Uncertainty

Sources of error include pre-processing operations associated with development of the land cover product as well as problems with pre-processing of data provided to the land cover process. Two phases of processing are necessary for generation of the MODIS Land Cover Product: data compositing and data analysis. In the compositing phase, the 32-day composited databases are assembled from MODIS Level 3 inputs. In the analysis phase, a year of 32-day composites are processed by the land cover and land-cover change algorithms to produce the quarterly output products.

3.1.5.1 Inputs

Both the Land Cover and Land Cover Type Change Parameters rely on a 1-km gridded database composited from MODIS Level 3 products. Inputs include (1) BRDF-corrected surface reflectances at 1000-m spatial resolution in the MODIS Land Bands (1-7) (MODIS Product MOD43); (2) spatial texture derived from Band 2 (near-infrared) at 250-m resolution; (3) vegetation index (MOD34); (4) weekly snow/ice cover (MOD33); (5) weekly land surface temperature at 1 km (MOD11); (6) directional reflectance information derived from the MODIS BRDF/Albedo Product (MOD43); and (7) ancillary terrain elevation information. These data are composited over a 32-day time period to produce a globally-consistent, multitemporal database on a 1-km grid as input to classification and change characterization algorithms (Strahler *et al.* 1995).

The ability to meet validation objectives may be influenced by changes in the input data. The input data can affect the land cover product because of the nature of the data or by affecting algorithm or process. This influence can be minimized if the algorithms are robust and promote multivariable, convergence-of-evidence approaches rather than relying on a single input parameter. At any rate, the nature and validity of input data must be monitored for its impact on land cover product validity.

3.1.5.2 Clouds

Persistent cloud cover will impede acquisition of high quality time trajectories of reflective or thermal data for use in characterizing land cover types. Even for a compositing period of 32 days, lack of cloud-free data will be a significant problem at some times of year in some regions, especially the humid tropics.

Cloud screening occurs in the production of several products that are input to the composited database. Therefore, the problem for the Land Cover Product should be one of missing rather than cloud-contaminated data. For isolated pixels and small regions of missing data, the solution is to provide temporally or spatially-interpolated values. When large areas of a monthly composite for a continent or large region show missing values, the appropriate strategy is to eliminate the composite from processing. The impact of loss of data due to clouds (and other causes) can be minimized by using algorithms that handle missing data well, which was one of the criteria for selecting an artificial neural network approach.

3.1.5.3 Registration

Misregistration may be another significant source of error. Since each MODIS measurement is geolocated, this problem amounts to uncertainty in true geolocation. Geolocation error is discussed in section 2.4.

A further consideration for geolocation and registration is continuity with ground location data. For field data to be effectively used for training and validation, transformations must be defined for coreferencing these data. This may require standardization of projection parameters such as datums and spheroids.

3.1.5.4 Gridding and Binning

Since Level-2 values are available, and in fact are used to characterize variance within grid cells, gridding will not introduce as much error or uncertainty as would occur if only Level-3 values were available (Kahn *et al.*, 1991). Multidate registration of Level-2 products to the Level-3 grid will be influenced by errors in geolocation. Excluding blunders, these are likely to be larger at larger scan angles. BRDF-fitted reflectance will be an improvement over selection by maximum value or by simply selecting against large view zenith angles. Output at 1-km resolution instead of the nominal 250-and 500-m resolutions for surface reflectance will reduce errors in spatial overlay. Moreover, if multidate registration proves problematic, it may be possible to cross-correlate images during the monthly compositing step and discard those that are poorly registered due to geolocation errors. This would require a significantly greater processing burden.

3.1.5.5 Topographic Data Error

Elevation is a key factor in geolocation of pixels. Elevation also must be accommodated in atmospheric correction, since it influences path length as a function of view angle. We may expect that both of these sources of error will be substantially corrected prior to compositing in the production of input products.

Pixel-scale topographic variations will, however, generate localized radiometric effects that influence surface reflectances. These effects include slope and aspect-related illumination differences, reflection from adjacent terrain, and subpixel shadowing that will generate BRDF effects. Proper correction for these effects requires accurate, high-resolution digital terrain data, which necessitates the use of the ancillary elevation data. These data may not be available at launch and are more likely a postlaunch consideration. The precise requirements for terrain data must be developed and reconciled within MODLand.

3.1.5.6 Data Dependencies

Since the Land Cover Product uses other MODIS products such as vegetation index, snow/ice cover and BRDF/Albedo products as well as a Land/Water Mask as inputs, its accuracy will depend partly on the accuracies of those products. However, it is not likely to be very sensitive to small errors in these input parameters. Based on the long history of successful land cover classification using remotely sensed data, we expect the spectral, spatial, temporal and directional signals to be quite robust in their information content for land cover, given that the instrument at least approaches its signal-to-noise ratio specifications. Because the classifier operates empirically, biases are not likely to be a problem as they might be for algorithms that produce quantitative geophysical parameters. Further, the compositing process that assembles the products will read quality flags and discard low-quality observations wherever possible.

For all data dependencies, it is necessary to track changes in inputs and their effect on algorithms, processing, flow, and product validity. Land cover and land-cover change are linked by the relationship of multitemporal characterization of land cover and multitemporal discrimination and description of land-cover change. This is manifest in the use of change detection to isolate multitemporal signal noise and change for validation of process and algorithm, as well as characterization of land cover types and optimization of field sampling.

3.1.5.7 Temporality

Error may be introduced when validation data are not temporally coincident with MODIS observation. This can especially be a problem with detecting and describing change. Although data compositing is prescribed, it will be necessary to define meaningful temporal generalizations of land cover based on successive observations. A further issue of temporality is assuring the temporal continuity of algorithms and processes.

3.1.5.8 Algorithm

Within the Land Cover Parameter, errors are generated when the classification algorithm selects the wrong class. With respect to a particular class, errors of omission occur when pixels of that class are assigned wrong labels; errors of commission occur when other pixels are wrongly assigned the label of the class considered. These errors occur when the signal of a pixel is ambiguous, perhaps as a result of spectral mixing, or when the signal is produced by a cover type that is not accounted for in the training process. These errors are a normal part of the classification process. They can be minimized, but not avoided entirely. Although they cannot be identified on a pixel-by-pixel basis due to processing constraints, they can be characterized in a statistical sense.

3.1.5.9 Reference Data

Reference data include both test site data and ancillary land cover and other environmental data. The quality and availability of adequate training/validation data derived from field sites and existing maps and tabular data is the most limiting factor to land cover and land-cover change validation (Muchoney *et al.* 1996). The quality of reference map data is a function of their inherent locational and thematic accuracy, while their utility may be restricted by incompatible classification systems or time of creation and validation.

Accuracy assessment of land cover products depends on the type and accuracy of reference (“truth”) data comprising field observation, remote sensing and collateral data sources. For individual test sites, the utility and quality and of the ground truth will be variable. Data utility is influenced by the classification system that is applied or parameters that are derived for a site. Site data utility is also a function of the source data since this impacts discrimination of features; what is observable using TM does not necessarily translate (at least directly) into comparable MODIS-discernible features. Source data also influence the minimum mapping unit and dimension.

The factor which primarily affects the quality of reference data is the underlying accuracy of the ground truth classification which may not be known. The time difference for the source data used in developing a reference dataset, the reference (presumably field-based) date and the MODIS acquisition dates impact both utility and accuracy of test site data.

Because factors relevant to validation vary considerably from test site to test site, validation will require assessment of the utility and accuracy of data available at each test site and most probably reworking site data to extract information specifically useful to land cover and land cover change. This argues strongly for the development of a high-resolution reference dataset that might be derived from other remote sensing sources, standardization of classification subunits or parameters, standardization of procedures for deriving classification subunits and parameters and development of a global sampling scheme and associated database.

High spatial resolution imagery will be available from a number of sources including ASTER, which will be on the EOS-AM platform with MODIS, the Landsat-7 ETM instrument, due for launch at about the time of the EOS-AM platform and destined for a near-simultaneous orbit and Lewis. With high spatial resolution data available, spatial heterogeneity of the test sites and the classes they contain can be further characterized and monitored. As a continuing data source, these instruments will also allow updating of land cover ground truth at test sites through the EOS era. Use of collateral remote sensing datasets such as TM and Lewis provide for a number of additional benefits of redundancy and complementarity that can be derived using data integration and data fusion techniques.

3.2 Practical Considerations

3.2.1 Numerical Computation Considerations

We do not anticipate problems with numerical stability and/or round-off errors.

3.2.2 Programming/Procedural Considerations

Two phases of processing are necessary for the MODIS Land Cover Product: data compositing and data analysis. In the compositing phase, the 32-day composited databases are assembled from MODIS Level 3 inputs. In the analysis phase, twelve 32-day composites are processed by the land cover algorithm, and 24 composites are processed by the land-cover change algorithm to produce the quarterly output products.

Allowing for quality flags, average daily volume will be approximately 0.51 GB. The processing power required is about 28 MFLOPS in order to assemble the 32-day database in one week. Obviously, this figure will increase if a more timely execution is required by the DAAC. At any rate, this power is about that of a mid-range engineering workstation.

Based on Version 1 software runs, generation of a single composite database for the terrestrial surface of the globe will require approximately 41.3 MFLOPS of CPU in the production environment. The corresponding output database will be approximately 5.34 GB in size. Similarly, in order to generate the quarterly products, 184.7 MFLOPS of CPU will be required to produce the 5.62 GB output databases. For both processes, these CPU estimates assume a 32-day period for completion.

3.2.3 Postlaunch Validation

3.2.3.1 Land Cover Parameter

Proper validation of a global dataset is not a simple task. Whereas validation of a biophysical parameter might entail developing a quantitative estimate or sense for the physical meaning of the parameter under consideration (Kahn *et al.*, 1991), land cover validation provides an indication or estimate of confidence that a pixel or segment has been correctly labeled as to a thematic class. Therefore, validity is dependent on how we define land cover classes. If the objective is to place a bounded estimate on the global per-pixel accuracy of the classification, then a formal sample design, based on a random, random-stratified, or systematic spatial sample, is required (Cochran, 1977). Such a sample requires obtaining reference data at many locations on the globe. The cost of acquiring such knowledge is therefore prohibitive, given the postlaunch resources for validation that we anticipate being available to the MODIS Land Team. Instead, we must turn to the test sites for which we have high-resolution land cover information available. Because the test sites are a biased sample, accuracy statistics derived at test sites cannot be regarded as proper statements of global accuracy. However, if the test sites are reasonably representative of their region as is planned, test site statistics can at least point to weaknesses and strengths in the dataset and allow users to anticipate how errors might impact their own research.

There are several approaches to the selection of pixels for comparison. First, accuracies may be reported by comparing the results obtained by the classifier in back-classifying training sites. Typically, this method is used to benchmark relative accuracy of classifiers rather than to establish a practical standard of accuracy. Given the nature of the algorithms, they back-classify (reclassify) training data to accuracies approaching 100 percent. Therefore, it is not useful to use this jackknife approach to assess thematic accuracy. Second, a set of test samples that is separate from training samples may be classified. Note that for convenience in processing, training pixels are sometimes included in this set when they comprise only a small proportion of the total test pixels. If these samples are selected according to a proper sample design, accuracies obtained by this method can be used to establish overall and per-class classification accuracy for the domain sampled (Green *et al.*, 1993). They may also be used to place bounds on areal estimates of coverage by class within the domain (Cochran, 1977).

Accuracy assessment has progressed through four development epochs over the last 25 years. The present stage may be described as the age of the error (confusion) matrix (Congalton, 1994). Classification accuracy is described using tables that document errors of commission and omission by cross-tabulating per-pixel labels output by the classifier with labels obtained from ground truth mapping or by classification of higher-resolution imagery (Story and Congalton, 1986). The kappa coefficient (Cohen, 1960) has become a standard statistic to evaluate overall classification accuracy, providing a more realistic estimation than a simple percentage agreement value. The kappa coefficient considers all cells in the confusion matrix, providing a correction for the proportion of chance agreement between the reference and test data sets (Rosenfield and Fitzpatrick-Lins, 1986). A *Z*-statistic can also be used as a pair-wise test of significance between two techniques based on the error matrices at specified probability levels (Congalton *et al.*, 1983). However, it has been found that kappa overestimates the proportion of chance agreement and consequently underestimates overall accuracy (Foody, 1992). Ma and Redmond (1995) present an alternative statistic for assessment of overall classification accuracy, the tau coefficient. This statistic is based on a *priori* probabilities of class membership rather than the *a posteriori* probabilities that are the basis for kappa. Tau is reported to better adjust percentage agreement to compensate for chance agreement, and to be easier to calculate and interpret. As with kappa, pair-wise tests of significance may be performed.

The confusion-table approach to accuracy assessment operates on the paradigm that each sample can be properly labeled into a single class, both by the classifier and by the process that establishes the ground reference (truth) data. It should be recognized that classification accuracy assessment may contain either conservative or optimistic bias. Simple interpretation of confusion matrices and related statistics without consideration of these error sources in the reference data may generate misleading conclusions (Verbyla and Hammond, 1995; Hammond and Verbyla, 1996). Sources of conservative bias in accuracy assessment (*i.e.* factors that reduce observed accuracies) include registration errors between reference and test data sets, use of a minimum mapping unit that is larger than the size of pixels in the classified image (Verbyla and Hammond, 1995), and the assumption that the reference data are perfectly correct (Congalton and Biging, 1992; Congalton and Green, 1993). Sources of optimistic bias (*i.e.* factors that increase observed accuracies) include sampling from training sites, non-independence of reference and training data and sampling from homogenous blocks of pixels (Hammond and Verbyla, 1996).

Accuracy statements about the product clearly depend on the accuracy of the ground truth. For individual test sites, the quality of the ground truth will be variable. Factors affecting the quality of the ground truth include (1) the underlying accuracy of the ground truth classification; (2) the units of land cover classification at the test site and their correspondence with those of the Land Cover Parameter; and (3) the difference in time between the acquisition of ground truth data and the remotely sensed data that are classified. Because these factors will vary from test site to test site, validation will require an individualized assessment of the characteristics of the product within ecoregions.

For each test site we will provide (1) class accuracy cross-tabulations resulting from a proper sample design; (2) a set of probabilistic statements of accuracy for the classification within the test site (3) a set of within-class error variances derived from the cross-tabulations that can be used to set standard errors on areal aggregations; and (4) a discussion and interpretation of accuracy issues and areal aggregation statistics that is geared to applications of the land cover parameter.

This validation strategy is similar to that adopted by the IGBP-DIS Land Cover Validation Working Group for the IGBP-DIS Global Land Cover Database (IGBP-DIS, 1995; see section 2.3.2). They plan analyses of classification accuracy and other characterization activities at a network of confidence sites (see section 3.1.3) similar to the test sites we plan to use in validation of the Land Cover and Land-Cover Change Parameters. As a pilot for technique development, we plan to participate in this effort.

The IGBP-DIS plan also includes a core sample of global land covers, selected by a formal sample design. At these sample sites, high resolution imagery (TM or SPOT) will be photointerpreted to validate the label of each pixel selected for sampling. The demands of this approach for high-resolution imagery as well as trained photointerpreters will require a separate funding initiative that the IGBP-DIS hopes to secure in the near future. Since these activities are planned for the prelaunch and early postlaunch periods, the IGBP-DIS validation activity will serve as a pathfinder for validation of the MODIS Land Cover Parameter.

Another important factor in test site analysis for product validation is that high spatial resolution imagery will be potentially obtainable from two sources: ASTER, which will be on the EOS-AM platform with MODIS, and the Landsat-7 ETM instrument, due for launch at about the time of the EOS-AM platform and destined for a near-simultaneous orbit. With high spatial resolution data available, spatial heterogeneity of the test sites and the classes they contain can be readily characterized and monitored. Further, as a continuing data source, these instruments will allow updating of land cover ground truth at test sites through the EOS era.

For IDS investigators using the Land Cover Product, quantitative accuracy assessment is only the first step in the analysis of how errors in the classification may affect their own applications. Through direct contacts with IDS investigators, we plan to explore and jointly develop validation methods for specific applications. In this way, we can ensure that the Land Cover Product meets the broadest range of EOS needs.

3.2.3.2 Land-Cover Change Parameter

Validating land-cover change maps is a complex task since it requires the observation of land-cover characteristics before and after an area is affected by a process of change. Our global test site development initiative (see Section 3.1.3) is producing an *a priori* list of test sites for land-cover change validation based on a sampling, deforestation "fronts," ecotones and ecological gradients, record of current change processes and "hot spots" where human pressure is high and where it is likely that land-cover conversion will take place. In most areas, a time interval of several years is necessary to detect significant land-cover changes and to be able to characterize accurately change processes and impacts. The case of the African Sahel is exemplary: several authors have shown that the interannual climatic variability in this region is such that only time series longer than a decade would allow for detection of any secular trend in land-cover change (Tucker *et al.*, 1991; Hellden, 1991). In addition to developing a sampling scheme for validating land-cover change, we are increasing our efforts in site-level analysis of change processes using AVHRR, Landsat and MODIS-simulated data.

3.2.4 Quality Control and Diagnostics

Quality control follows the MODLand Quality Assurance (QA) Plan (Roy, 1996). The plan outlines run-time and post run-time QA procedures for MODLand standard products. The QA data fields consist of those mandated by ECS, those common to all MODLand products, and product-specific metadata. Run-time QA information is generated in the production environment, and is either spatially explicit (per-pixel) or global (per-tile) in scope.

Mandatory MODLAND QA is generated on a per-pixel basis. The 8-bit flag consists of 3 cloud-state bits, 1 bit describing product usefulness and up to 4 additional bits for product summary as specified by the science team member responsible for each product. For the MODIS Land Cover Product, these additional bits are a summary of the full set of QA statistics generated for the product.

Run time quality assurance data specific to the Land Cover Parameter primarily conveys confidence in the label of each grid cell as an overall quality flag, as generated by the neural network classifier. Other flags indicate secondary issues regarding the quality of the input database. Run time QA issues related to the Land-Cover Change Parameter are still in early research stages owing to the post-launch status of the data product.

Post run time QA is generated at the SCF and at the Land Data Operational Product Evaluation (LDOPE) facility, a centralized QA installation. The main role of the LDOPE is to carry out routine QA evaluation, while the SCF staff handles situations that require greater scientific expertise. When data fail any quality test at the DAAC or LDOPE (as defined by either ECS or the SCF), the SCF will be notified by the DAAC or LDOPE. At that point, SCF staff may elect to examine the data at the SCF, or in cooperation with LDOPE personnel. Results of post run time QA are then sent to the DAAC, where they are included as part of the mandatory ECS metadata.

3.2.5 Exception Handling

Exception handling, which covers data generated during infrequent events such as platform maneuvers, eclipses, and the like, will primarily be the responsibility of the input products. Thus, these events will produce missing data fields.

3.2.6 Data Dependencies

Since the Land Cover Product uses Level-2 and Level-3 products as inputs, its accuracy will depend partly on the accuracies of those products. However, it is not likely to be very sensitive to small errors in these input parameters. Based on the long history of successful land cover classification with remotely sensed data, we expect the spectral, spatial, temporal and directional signals to be quite robust in their information content for land cover, given that the instrument at least approaches its signal-to-noise ratio specifications. And, because the classifier operates empirically, biases are not likely to be a problem, as they might be for algorithms producing quantitative geophysical parameters. Further, the compositing process that assembles the products will read quality flags and discard low-quality observations wherever possible.

3.2.7 Output Products

Output from the Land cover Product will include the two parameters, Land cover and Land-Cover Change, encoded as categorical variables that are stored as byte data. Accompanying each byte will be approximately 15 bytes of attribute information which consist of entries such as quality flags and data-field characteristics. Assuming, then, that each parameter requires 16 bytes of data, the global data volume per parameter is about 8.8 GB, based on a figure of about 1.5×10^8 sq km of earth land surface.

Because both products require a temporal sequence of data, neither will be available immediately after launch. However, a provisional product (see section 2.3.2) will be available at launch for the Land Cover Parameter. We anticipate releasing the first Land Cover Parameter at 18 months after launch, revising the product quarterly thereafter. The first Land-Cover Change Parameter will be released 27 months after launch, based on two years of monthly composites.

The Land Cover GCM Product is especially tailored for global climate modeling at coarse resolutions. It will be provided on a quarterly basis and will contain the proportions and areal estimates associated with each land cover class within a 1/4° grid cell. The exact format of this product is still under development.

4. Constraints, Limitations, Assumptions

Constraints, limitations and assumptions are discussed in appropriate sections *ad seriatim* in the preceding text of this document. For example, both the Land Cover and Land-Cover Parameters require properly registered and resampled data that are cloud-screened and atmospherically corrected (sections 3.1.1.1.1, 3.1.5).

5. Additional Postlaunch Products

5.1 Relationship of Land Cover Products Proposed by the University of Maryland with Earlier Proposals

Two additional products have been proposed by the University of Maryland which are complementary to those proposed in the original ATBD namely:

i) A land cover change indicator product at a resolution of 250m. This product is designed to provide early warning that land cover change is occurring. The limited spectral characteristics of the 250m bands will limit identification of the type of change that is occurring but the high spatial resolution should substantially improve the timeliness by which the existence of change is flagged compared with data at 1 km or even 500m.

ii) A representation of global land cover by continuous fields rather than by classes. This product is designed to overcome the arbitrariness of abrupt distinctions between classes, by representing land cover in each pixel by a proportion of basic cover components such as percentage tree cover or percentage herbaceous cover.

5.2 A Land-Cover Change Alarm Product at a Resolution of 250m

5.2.1 Overview and Background

A series of simulation experiments using data from the Landsat Multispectral Scanner System indicate the effect of spatial resolution on the detection of change of a key biophysical variable namely the normalized difference vegetation index: this has been widely used in the classification of land cover (Townshend *et al.* 1988), (Townshend *et al.* 1990), (Townshend *et al.* 1995). The results of these analyses indicate that a very high proportion of changes, which primarily arise because of land cover conversions, occur at the finest spatial resolutions. It was therefore proposed that two 250m resolution bands should be included on MODIS (Townshend *et al.* 1988). The benefits of a 250m resolution over a 500m resolution as result of this work are indicated in figure 1. These results indicate that the use of data with a 250m resolution has the potential to be substantially more sensitive to land cover change than the 500m or 1 km bands. We intend to generate a global change detection product indicating the distribution of land cover change. It will be important to attempt to distinguish between actual land cover changes from other types of changes. The latter will include changes due to phenology, snow cover, and cloud. We will use cloud and snow products derived by other team members to avoid identification of spurious changes due to the latter and will use change detection methods designed to “filter-out” phenological changes.

It is proposed that the change detection product is used to contribute to an alarm system to indicate the areas of highest priority where fine resolution data should be collected. It is anticipated that Landsat 7 data will provide much better coverage than previously, since the space and ground segment will be capable of acquiring many more images per day. Nevertheless the instrument will inevitably remain a sampler in time and space compared with MODIS and decisions will have to be made about which areas to image and which not to in a given cycle. Knowledge of where land cover conversions are taking place will provide useful intelligence in developing the acquisition strategies since these areas must be of high priority for monitoring. We proposed to develop close liaison with the Landsat Science Team to assure the successful implementation of this approach.

5.2.2 Algorithm Description

The overall processing sequence is shown in figure 2. We will primarily exploit the multispectral and multitemporal information available in the two spectral bands with a resolution of 250m. Although there are only two bands available at this resolution they are found in two key spectral regions. The behavior of reflectances of these bands is rather well known and can be exploited to identify land cover change: for example ratios of the red and the near infrared are especially sensitive to green leaf properties and clearances in land cover are also often associated with increases in albedo, whereas inundation (unless below a dense vegetation canopy) typically greatly reduces reflectance especially in the near infrared.

In deriving the 250m products we will use gridded (level 2G) products of spectral reflectance for these two, where atmospheric corrections for molecular scattering, ozone, water vapor and tropospheric and atmospheric aerosols have been carried out. Generation of these products is already the responsibility of existing team members.

Data from individual dates will have to be formed into temporal composites so that the effects of clouds are minimized. Currently funded research of the proposed team member (Pathfinder AVHRR version-2) is investigating methods of compositing in addition to the conventional maximum likelihood procedure. The latter produces fields of vegetation indices, satisfactory for many purposes, but produces highly variable fields of individual bands. We are examining combining layers from individual dates, with minimum view angles, having flagged the presence of clouds using combined information from the reflective and thermal bands. We will investigate the use of MODIS cloud products for cloud screening.

It is anticipated on the basis of experience with AVHRR data that temporal composites of at least a month will be required to generate data sets in which cloud effects do not prevent effective use.

Two related approaches are proposed in using the spectral bands for change detection, a two-date procedure where data from only two composites are used and a multidate procedure where deviations from "normal" seasonal variation are used to identify change.

5.2.2.1 Two-date Procedure

To use the limited information available in the two spectral bands we need to use properties in which changes extrinsic to those of land cover change are minimized. The use of atmospherically corrected, calibrated spectral data will do much to assist this.

Our intention is to examine a number of procedures to detect changes between dates. As an example of one possible approach we note that two of the main changes in spectral properties associated with land cover change conversions are albedo and greenness. Changes in the normalized difference vegetation index and overall reflectance from the red and near infrared should provide a useful operational approach in the detection of land cover change.

Even allowing for the correction of atmospheric and calibration effects, many changes which occur within scenes are unrelated to actual land cover changes and hence it is necessary to identify a thresholds of spectral response beyond which "true" changes actually occur. One possibility is to use the minimum size of the thresholds VI_{th} and R_{th} needed to detect the occurrence of cover change. As figure 3 indicates six basic types of change can be conceptually identified, namely where VI_{th} or R_{th} will be exceeded during either an increase or decrease from time 1 to time 2 and where both thresholds are exceeded. It may be necessary to vary the size of the threshold according to the absolute size of the value of the vegetation index since the latter has been shown to vary non-linearly with some vegetation properties (Curran 1983).

Many of the changes which may be observed may simply arise due to phenological changes. Hence we will investigate the use of changes in values of the vegetation index and brightness in excess of the mean changes within a window of dimensions W surrounding each pixel. Thus we will apply a rule of the form:

$$Flag\ \ change\ \ IF \left[P - \sum_{i=1}^{Nw} p \right] > Th$$

Experiments will be carried to determine the size of W. Given that most significant changes in land cover will have small dimensions, it is not anticipated that a window size of more than 20-30 pixels will be required and it may be possible to use an even smaller window size thus reducing the size of the computations.

The result of this procedure will therefore be a data plane of substantial increases and decreases in greenness and brightness. These will be further characterized by the initial state of each pixel (for example whether an increase in greenness occurred for a bright or dark target). Given the limited spectral information available from the two spectral bands it will probably not be possible globally to assign a unique land cover conversion label to each type of trajectory within the VI-R space. However it is anticipated that through use of models of ecological succession and anthropogenic disturbance (Running, *et al.* 1994) and through empirical analysis of both Landsat and AVHRR data it will be possible to build up regional typologies of characteristic land cover changes associated with these trajectories. Figure 4 shows an example of the changes in albedo and greenness corresponding to typical land cover changes in tropical rain forests.

To develop the regional models and typologies the proposed team member and current collaborators in other funded projects will use their substantial remote sensing and field experience from many regions of the world. Collaborators are involved in a number of projects including

i) tropical deforestation and regrowth in the Pan-Amazon and Central Africa,

ii) savanna dynamics in east and Sahelian Africa,

iii) boreal forest dynamics in Canada,

iv) desert boundary changes in central Asia and northern Africa,

v) temperate forest and agricultural change in the United States,

vi) coastal wetland change in the eastern United States.

This knowledge will be supplemented, taking advantage of other validation activities.

5.2.2.2 Multi-Date Procedure

The procedures used in the multidate approach will follow fundamentally the same approach as that used in the dual-date approach, but the threshold of changes will be also carried out relative to the annual phenological cycle of greenness displayed by vegetation indices when plotted against time (*e.g.*, Justice *et al.* 1985). Once more than a year's data have been obtained we will then use deviations from the seasonal vegetation index temporal profiles to assess whether the threshold has been exceeded indicating a land cover change. Consideration of the observed pixel changes relative to changes in surrounding pixel values should also be included in the decision rule, as proposed for the dual date approach, since otherwise many of the flagged changes in land cover will relate merely to local and regional phenological variations arising because of inter-annual variability in factors such as rainfall.

It is currently unclear whether a useful annual profile of changes in the individual spectral bands can be used in the same way because of the impacts of bi-directional effects. The creation of improved data sets at 1 km resolution as part of the AVHRR Land

Pathfinder Version 2 should allow this to be determined before the launch of EOS-AM. If such seasonal variability cannot be used will then rely on the prior brightness to help assess whether change has occurred in similar way to that described in above. Overall we would anticipate that more consistent land cover change products will be created by this multidate approach.

Initially the land cover change product will predominantly be an indicator that certain types of change are likely to have occurred. Labeling of the actual changes will be assisted by the application of models of land cover change. These will be based on the actual land cover present before the change since this will substantially reduce the range of changes that could occur. Also we intend to use the flags to influence the acquisition of Landsat data and analyze the to allow more specific labeling of actual conversions to be provided.

5.2.2.3 Use of Changes in Local Spatial Variability

One issue which will need to be considered is the impact of misregistration on the change detection product. The current expected 100m 2 x RMS error between dates is at the margins of acceptability when working at 250m pixel size, because of the introduction of differences solely as a result of misregistration (Townshend *et al.* 1992). We will examine the consequences of these errors on our proposed procedures. It is unclear whether local correlation methods to improve registration will be successful because the images to be used will be composites derived from multiple dates We are therefore currently investigating the use of change detection methods based on changes in variability of the image. This approach is justified by the piecemeal nature of most change in land cover leading to increased variability when it occurs. Such procedures can be used with windows of sufficient size that the expected misregistration of the images has a small impact on change detection. Preliminary experiments have indicated the promise of this procedure.

5.3. Representation of Land Cover by Continuous Fields Using MODIS Data.

5.3.1 Overview and Background

As part of our NASA funded Global Land Cover Classification project we are using training sites developed from analysis of existing ground based maps and fine resolution satellite data to produce land cover classifications of the globe. This is the land cover class product which is published on the ISLSCP CD-ROM. Using our existing methods and training data we are currently generating a product at 8 km using the AVHRR Land Pathfinder Version 1 data set. This will form a contribution to the ISLSCP Version 2 activity. We further intend to generate a land cover product at 1 km using the USGS/NASA/ NOAA/IGBP-global AVHRR data set using current funding from NASA. This is offered as a contribution to an at-launch MODIS product and we will work to convert this or other at-launch data sets of land cover into forms needed to support other MODIS products.

Instead of the conventional land cover classifications, recent work at the University of Maryland has shown great promise for the creation of products of continuous attributes of vegetation components such as the proportions of tree cover, bare ground and the herbaceous layer by use of methods such as empirical regressions, decision-tree procedures and mixture modeling (Defries *et al.* 1995). Conventionally methods such as mixture modeling have relied on unmixing using multispectral data, but this recent work has also exploited the potential of multitemporal sequences (Hanan *et al.* 1991). These methods are presently under active development through another NASA grant and a NSF Grand Challenge grant. Preliminary results are shown in figures 8 and 9 when applied to coarser resolution AVHRR data at 8 km and 1 degree resolution: they indicate the considerable potential of this approach (DeFries *et al.* 1995). As with many global data sets validation remains difficult (see section 5.7) but comparisons with results obtained from fine resolution data indicate that the predicted forest cover was within 20% of the actual forest cover for 90% of the pixels (Defries *et al.* in press). We propose to apply these methods initially to the version-1 1 km AVHRR global data set and subsequently to version 2 data sets as they become available and to make these representations of vegetation properties available as at-launch products. Currently it seems feasible to be confident of producing global products of tree cover and proportion of bare ground and additional products may be possible

5.3.2 Algorithm Description

Given the success of the prototype data sets of land cover components, it is proposed that these methods are applied to MODIS data to create post-launch products of land cover as represented by continuous fields rather than traditional classes. Such data should yield considerably improved products compared with those obtained with AVHRR data, because of the removal of atmospheric effects and because of the higher spectral dimensionality of the MODIS data sets compared with the AVHRR. We therefore propose to use these techniques based on existing research to create fields of vegetation properties. Because we have no direct analogs of MODIS data with its combined multitemporal and multispectral richness it is difficult to be certain of exactly what properties of the vegetation can be successfully characterized. For example in mixture modeling it is unclear which end-members will be successfully extracted and what are the resultant continuous properties we might hope to extract, but we should at least improve the reliability of those extracted already. In the case of the 250m bands we would of course largely have to exploit the multitemporal approaches alone.

We are currently using two methods to derive continuous fields of vegetation properties with AVHRR data: linear mixture modeling and empirical estimation from multitemporal data. We expect to apply these methods to MODIS data from the 250 m bands, with superior results due to improved atmospheric correction and higher spatial and spectral resolution.

5.3.2.1 Linear Mixture Modeling

Linear mixture modeling is based on the assumption that spectral values in a pixel contain information about the proportion of each component present within that pixel (Settle and Drake 1993, Shimabukuro and Smith 1991, Smith, *et al.* 1990) The contribution of each component to the spectral value is assumed to be linearly proportional to the aerial coverage, hence:

$$r_i = \sum_{j=1}^{n} \left(a_{ij} x_j\right) + e_i$$

where r_i = mean spectral reflectance for the ith spectral band of a pixel containing one or more components; a_{ij} = spectral reflectance of the *j*th component in the pixel for the *i*th spectral band; x_j = proportion value of the *j*th component in the pixel; e_i = error term for the *i*th spectral band; $j = 1, 2, \ldots, n$ (number of components assumed for the problem); and $i = 1, 2, \ldots, m$ (number of spectral bands from the sensor system used for the problem).

The method is sensitive to the values selected as “endmember” values (a_{ij}). End members are pure pixels with 100 percent coverage of vegetation with the respective characteristic. We will use coregistered high resolution Landsat data to identify locations of pure pixels. To date, we have used locations known through field studies to have 100 percent coverage as end members.

We have used multitemporal spectral data from the AVHRR Pathfinder Land data set, namely mean annual values for red and near infrared reflectances, in the mixture modeling equation to obtain estimates of proportions for three components: woody vegetation, herbaceous cover, and bare ground (DeFries, *et al.* 1996). Each component is thus identified by its phenological variability as well as its spectral response at a single data.

5.3.2.2 Empirical Estimation of Proportional Cover

An alternative method to estimate proportional woody cover involves empirical linear relationships between percent forest cover derived from coregistered high resolution Landsat data and metrics derived from AVHRR multitemporal data. The use of multitemporal data improves estimates substantially over data from a single month (DeFries, *et al.* in press). In summary, the method involves 1) derivation of proportional cover at a coarser AVHRR resolution from coregistered, classified Landsat scenes; 2) stepwise linear regression to derive a relationship between percent woody cover as the independent variable and multitemporal metrics, such as mean annual NDVI and difference between annual maximum and minimum NDVI, as dependent variables; and 3) extrapolation of the relationship over a larger area to predict percent woody cover. We have applied this method to central Africa and are currently extending it to the global scale.

5.3.2.3 Proportional Cover from MODIS 250 m Bands

We will test both methods with MODIS data simulated from TM data to determine which gives more accurate results. For the post-launch product, calibration data to derive the empirical relationships or, in the case of mixture modeling, the endmember values will be selected from the global network of high resolution Landsat data that has been assembled at the University of Maryland, as well as other test sites as they become available through the Global Land Cover Test Sites project and other activities. The Landsat data will be coregistered with MODIS data through the use of control points. Sites known to be homogeneous at a 250 m resolution will be selected as end members for mixture modeling. Proportional cover will be derived from coregistered, classified Landsat data. The proportional covers will be used for calibrating the empirical relationships and for validating the results.

We will first derive continuous fields indicating aerial proportions of woody vegetation, herbaceous cover, and bare ground. We also plan to use similar methods to derive continuous fields of aerial proportions of broadleaf and needleleaf vegetation, as well as deciduous and evergreen vegetation. If satisfactory results are not obtained for continuous fields of broadleaf vs. needleleaf vegetation and deciduous vs. evergreen vegetation, we will base these proportions on the 1 km land cover classification product as estimates for these vegetation properties. The methods for deriving continuous fields depends on multitemporal data from an annual phenological cycle, so we expect to generate the products following one year after launch.

5.3.2.4 Validation of Continuous Fields

Estimates of proportional covers will be validated with coregistered Landsat data. We currently have a globally distributed network of over 160 scenes from which to derive validation data. For each scene we have ancillary sources or field knowledge to aid the interpretation. Proportional covers for validation will be obtained through degradation of the classified Landsat scenes. In addition, IGBP test sites and Global Land Cover Test Sites will also be used for validation as they become available.

6. References Cited

Abuelgasim, A.A., S. Gopal, J.R. Irons and A.H. Strahler. 1996. Classification of ASAS multiangle and multispectral measurements using artificial neural networks. *Remote Sensing of Environment* 57:79-87.

Abuelgasim, A. and S. Gopal. 1994. Classification of multiangle and multispectral ASAS data using a hybrid neural network model. *Proceedings of the 1994 International Geoscience and Remote Sensing Symposium*, Aug. 8-12, 1994, Pasadena, CA, 3:1670-1675.

Barker, J.L., B.L. Markham and J. Burelbach. 1992. *Procedure for Spatial Simulation of MODIS data from TM imagery*. NASA/GSFC.

Barnsley, M., K. Morris and A. Reid. 1990. Preliminary analysis of a multiple-view-angle image data set. *Proc. NERC 1989 Airborne Remote Sensing Campaign Symp.*, Keyworth, England, 18-19 Dec., 1990, pp. 49-68.

Barnsley, M.J. 1995. The potential for land cover mapping using multiple-view-angle (MVA) image data. *Remote Sensing of Environment*: submitted.

Belward, A. and T.R. Loveland. 1995. The IGBP 1km Land Cover Project. In *Proceedings of the 21st Annual Conference of the Remote Sensing Society*, Southampton, UK, pp. 1099-1106.

Belward, A. (ed.). 1996. *The IGBP-DIS Global 1-km Land Cover Data Set DISCover*, IGBP-DIS Working Paper, in press.

Benediktsson, J.A., P.H. Swain and O.K. Ersoy. 1990. Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Trans. Geoscience and Remote Sensing* 28(4) 540-552.

Benediktsson, J.A., P.H. Swain and O.K. Ersoy. 1993. Conjugate-gradient neural networks in classification of multisource and very-high-dimensional remote sensing data. *International Journal of Remote Sensing* 14: 2883-2903.

Berk, A., L.S. Bernstein, and D. C. Robertson. 1989. *MODTRAN: A model resolution model for LOWTRAN 7*. Rep. AFGL-TR-89-0122, Air Force Geophysics Laboratory, Bedford, MA.

Bicheron, P., M. Leroy and O. Hautecoeur. 1996. Enhanced discrimination of boreal forest covers using directional signatures measured by the airborne POLDER instrument. *Proceedings of the International Workshop on Multiangular Remote Sensing: Measurements, Models and Applications,* September 13-18, 1996, Beijing, China (forthcoming).

Birdle, J.S. 1990. Probabilistic interpretations of feedforward classification network outputs with relationships to statistical pattern recognition. In, F. Fogelman Soulie and J. Herault, eds., *Neurocomputing*. Springer-Verlag.

Bischof, H., W. Schneider and A.J. Prinz. 1992. Multispectral classification of Landsat-images using neural networks. *IEEE Trans. Geoscience and Remote Sensing* 30:482-490.

Borak, J.S. and A.H. Strahler. 1996. Feature selection using decision trees -- an application for the MODIS land cover algorithm. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Lincoln, Nebraska, May 27-31, Vol. 1, pp. 243-245.

Borak, J.S. and A.H. Strahler. Feature selection and land cover classification of a MODIS-like data set for a semiarid environment (in preparation).

Breiman, L., J. H. Friedman, R. A. Olsen and C. J. Stone. 1984. *Classification and Regression Trees*, Wadsworth, Inc., Belmont, CA.

Brodley, C.E. and M.A. Friedl.1996. Improving automated land cover mapping by identifying and eliminating mislabeled observations from training data. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Vol. 2, pp. 1379-1381

Brodley, C.E., M.A. Friedl and A. H. Strahler. 1996. New approaches to classification in remote sensing using homogeneous and hybrid decision trees to map land cover. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium*, Lincoln, Nebraska, May 27-31, Vol. 1, pp. 532-534.

Brodley, C.E. and P.E. Utgoff. 1992. *Multivariate versus Univariate Decision Trees*. COINS Technical Report 92-8, 15 pp.

Brodley, C.E. and P.E. Utgoff. 1995.Multivariate decision trees. *Machine Learning* 19(1): 45.

Brown, D.E., V. Corruble and C.L. Pittard. 1993. A comparison of decision tree classifiers with backpropogation neural networks for multimode classification problems. *Pattern Recognition* 26: 953-961.

Byrne, G.F., P.F. Crapper and K.K. Mayo, 1980. Monitoring of Land-Cover Change by Principal Components Analysis of Multitemporal Landsat Data. *Remote Sensing of Environment* 10: 175-184.

Carpenter, G.A. 1989. Neural network models for pattern recognition and associative memory. *Neural Networks* 2:243-257.

Carpenter, G, M. Gjaja, S. Gopal and C. Woodcock. 1996. ART networks in remote sensing. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium*, Lincoln, Nebraska, May 27-31, Vol. 1, pp. 529-531.

Carpenter, G.A., S. Grossberg. 1987a. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Comp. Vision, Graphics and Image Processing* 37: 54-115.

Carpenter, G.A., S. Grossberg. 1987b. ART 2: Stable self-organization of pattern recognition codes for analog input patterns. *Appl. Optics* 26:4919-4930.

Carpenter, G.A., S. Grossberg and J.H. Reynolds. 1991a. ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks* 4: 566-588.

Carpenter, G.A., S. Grossberg and D.B. Rosen. 1991b. Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks* 4:759-771.

Carpenter, G.A., S. Grossberg, N. Markuzon, J.H. Reynolds and D.B. Rosen. 1992. Fuzzy ART: A neural network architecture for incremental supervised learning of analog multidimensional maps. IEEE Trans. *Neural Networks* 3:698-713.

Cochran, W.G. 1977. Sampling Techniques. John Wiley and Sons, New York.

Cohen, J. 1960: A coefficient of agreement for nominal scales. Educational and Psychological Measurement 10(2): 37-40.

Congalton, R. G. 1994. Accuracy assessment of remotely sensed data: future needs and directions. *Proceedings 12th Pecora Symposium*, ASPRS, Maryland: Bethseda, pp. 385-388.

Collins, J.B and C.E. Woodcock. 1994. Change detection using the Gramm-Schmidt transformation applied to mapping forest mortality. *Remote Sensing of Environment* 50:267-279.

Colwell, J. E. and F.P. Weber. 1981. Forest change detection, *Proc. 15th Int. Symp. on Remote Sensing of Environment*, Environ. Res. Inst. Of Mich., Ann Arbor, pp. 839-852.

Congalton, R.G. and G.S. Biging. 1992. A pilot study evaluating ground reference data collection efforts for use in forest inventory. *Photogrammetric Engineering and Remote Sensing* 58(12): 1669-1671.

Congalton, R.G. and K. Green. 1993. A practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering and Remote Sensing* 59(5): 641-644

Congalton, R.G., R.G. Oderwald and R.A. Mead. 1983. Assessing landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric Engineering and Remote Sensing* 49(12): 1671-1678.

Curran, P.J. 1983. Estimating green LAI from multispectral aerial photography. *Photograpmmetric Engineering and Remote Sensing* 49: 1709-1720.

Decatur, S.E. 1989. Application of neural networks to terrain classification. *Proc. 1989 Int. Joint Conf. on Neural Networks*, I-284.

DeFries, R.S., C.B. Field, F. Inez, C.O. Justice, P.A. Matson, M. Matthews, H.A. Mooney, C.S. Potter, K. Prentice, P.J. Sellers, J. Townshend, C.J. Tucker, S.L. Ustin, and P.M. Vitousek. 1995. Mapping the land surface for global atmosphere-biosphere models: toward continous distributions of vegetation's functional properties. *Journal of Geophysical Research* 100(D10): 20867-20882.

DeFries, R.S. and J. R.G. Townshend. 1994. NDVI-derived land cover classifications at a global scale. *International Journal of Remote Sensing* 15(17): 3567-3586.

DeFries, R., M. Hansen, M. Steininger, R. Dubayah, R. Sohlberg, and J. Townshend. Subpixel forest cover in Central Africa from multisensor, multitemporal data. *Remote Sensing of Environment.* In press.

DeFries, R., M. Hansen, and J. Townshend.1996. Proportional estimation of land cover characteristics from satellite data. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Lincoln, Nebraska, May 27-31, Vol. 1 , pp. 535-537.

DeFries, R., M. Hanson and J. Townshend. 1995. Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. *Remote Sensing Environment*. 54: 209-222.

Dorman, J.L. and P.J. Sellers. 1989. A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (SiB). *J. Appl. Meteorol.* 28: 833-855.

Dreyer, P. 1993. Classification of land cover using optimized neural nets on SPOT data. *Photogrammetric Engineering in Remote Sensing* 5: 617-621.

Duvernoy, J. and J. Leger, 1980. Karhunen-Loeve analysis of multispectral data from landscapes. *Optics Communications* 32(1): 39-44.

Fischer, M. M. and S. Gopal. 1992. Neural network models and interregional telephone traffic: Comparative performance between multilayer feedforward networks and the conventional spatial interaction model. *WSG Discussion Paper* 27/92, Department of Economic and Social Geography, Vienna University of Economics and Business Administration.

Fitzgerald, R.W. and B.G. Lees. 1994. Assessing the classification accuracy of multisource remote sensing data. *Remote Sensing of Environment* 47: 362-368.

Foody, G. M. 1992. On the compensation for chance agreement in image classification accuracy assessment. *Photogrammetric Engineering and Remote Sensing* 58(10): 1459-1460.

Gopal, S. and M. Fischer. 1996. A comparison of three neural network classifiers for remote sensing classification. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium*, Vol. 1, pp. 787-789 .

Gopal, S. and C. Woodcock. 1996. Remote sensing of forest change using artificial neural networks. *IEEE Transactions on Geoscience and Remote Sensing,* in press.

Gopal, S., C. Woodcock and A.H. Strahler. 1996. Fuzzy ARTMAP classification of global land cover from AVHRR data set. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium,* Lincoln, Nebraska, May 27-31, Vol. 1, pp. 538-540.

Gopal, S., D.M. Sklarew and E. Lambin. 1994. Fuzzy-neural networks in multitemporal classification of landcover change in the Sahel. *New Tools for Spatial Analysis, Proc. of the Workshop, Lisbon, 18 to 20 November 1993,* Office for Official Publ. Eur. Communities, Luxembourg, pp. 69-81.

? Gopal, S. and C. E. Woodcock. 1995. Remote sensing of forest structure and change using artificial neural networks. *Proc. ASPRS/ACSM Ann. Conf.,* Feb. 27-Mar. 2, 1995, Charlotte, NC.

Green, E.J., W.E. Strawderman and T. M. Airola. 1993. Assessing classification probabilities for thematic maps. *Photogrammetric Engineering and Remote Sensing* 59: 635-639.

Hammond, T.O. and D.L. Verbyla. 1996. Optimistic bias in classification accuracy assessment. *International Journal of Remote Sensing* 17(6): 1261-1266.

Hanan, N. P., S. D. Prince and P. H. Y. Hiernaux.1991. Spectral modelling of multi-component landscapes in the Sahel. *International Journal of Remote Sensing* 1243-1258.

Hansen, M., R. Dubayah and R. DeFries. 1996. Classification trees: An alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, 17: 1075-1081.

Hara, Y., R.G. Atkins, S.H. Yueh, R.T. Shin, and J.A. Kong. 1994. Application of neural networks to radar image classification. *IEEE Trans. Geoscience and Remote Sensing* 2: 100-11.

Heermann, P.D. and N. Khazenie. 1992. Classification of multispectral remote sensing data using a back-propagation neural network. *IEEE Trans. Geoscience and Remote Sensing* 30: 81-88.

Hellden, U. 1991. Desertification: time for an assessment? Ambio 20: 372-383.

Hellden, U. and L. Eklundh. 1988. *National Drought Impact Monitoring, A NOAA NDVI and Precipitation Data Study of Ethiopia,* Lund Studies in Geography, Ser. C., No. 15, Lund Univ. Press, Sweden, 55 pp.

Henderson-Sellers, A. and A. J. Pitman. 1992. Land-surface schemes for future climate models: Specification, aggregation and heterogeneity. *J. Geophys. Res.* 97: 2678-2696.

Henderson-Sellers, A., M F. Wilson and G. Thomas. 1985. The effect of spatial resolution on archives of land cover type. *Climatic Change* 7: 391-402.

Hepner, G. F., T. Logan, N. Ritter and N. Bryant. 1990. Artificial neural network classification using a minimal training set: Comparison to conventional supervised classification. *Photogrammetric Engineering and Remote Sensing* 56(4): 469-473.

Holben, Brent N. 1986. Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing* 7(11): 1417-1434.

IGBP. 1992a. *Requirements for Terrestrial Biospheric Data for IGBP Core Projects.* IGBP-DIS Working Paper #2 (June 1992), IGBP-DIS, Universite de Paris, Paris, France, 24 pp. + appendixes.

IGBP. 1992b. *Improved Global Data for Land Applications.* IGBP Global Change Report No. 20, IGBP, Int. Council of Sci. Unions, Stockholm, Sweden, 85 pp.

IGBP-DIS. 1995. *The IGBP-DIS Global 1 Km Land Cover Data Set: A Validation Strategy,* A. Belward, ed., IGBP-DIS, Universite de Paris, Paris, France.

Ingebritsen, S. E. and R. J. P. Lyon. 1985. Principal components analysis of multitemporal image pairs. *International Journal of Remote Sensing* 6:687-696.

Irons, J.R., K.J. Ranson, D.L. Williams, R.R. Irish and F.G. Heugel. 1991. An off-nadir pointing imaging spectroradiometer for terrestrial ecosystem studies. *IEEE Trans. Geoscience and Remote Sensing* 29:66-74.

Jackson, R.D. 1983. Spectral indices in n-space. *Remote Sensing of Environment* 13: 409-421.

Jupp, D.L.B., A.H. Strahler and C.E. Woodcock. 1988. Autocorrelation and regularization in digital images I: Basic theory. *IEEE Trans. Geoscience Remote Sensing* 26:463-473.

Jupp, D.L. B., A.H. Strahler and C.E. Woodcock. 1989. Autocorrelation and regularization in digital images II: Simple image models. *IEEE Trans. Geoscience Remote Sensing* 27:247-258.

Justice, C.O., J.R. G. Townshend, B.N. Holben and C.J. Tucker.1985. Analysis of the phenology of global vegetation using meteorological satellite data. *International Journal of Remote Sensing* 6(8): 1271-1381.

Kahn, R. R.D. Haskins, J.E. Knighton, A.Pursch and S. Granger-Gallegos. 1991. Validating a large geophysical dataset: Experiences with satellite-derived cloud parameters. *Proceedings of the 23rd Symposium on Interface, Computing Science and Statistics*, pp. 133-140.

Kanellopoulos, I., A. Varfis, G.G. Wilkinson, and J. Megier. 1992. Land-cover discrimination in SPOT HRV imagery using an artificial neural network 20-class experiment. *International Journal of Remote Sensing* 13(5): 917-924.

Kauth, R.J. and G.S. Thomas. 1976. The tasseled cap—A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Symposium Proceedings on Machine Processing of Remotely Sensed Data*, June 29-July 1, 1976, West Lafayette, IN, pp. 41-51.

Key, J., J.A. Maslanik, and A.J. Schweiger. 1989. Classification of merged AVHRR and SMMR arctic data with neural networks. *Photogrammetric Engineering and Remote Sensing* 55(9): 1331-1338.

Kiang, R.K. 1992. Classification of remotely sensed data using OCR-inspired neural network techniques. *Proc. Int. Geoscience Remote Sensing Symp. 1992*, May 26-29, 1992, Houston, TX, 2:1081-84.

King, M.D., Y.J. Kaufman, W.P. Menzel, and D. Tanre. 1992. Remote sensing of cloud, aerosol and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS). *IEEE Trans. Geoscience and Remote Sensing* 30: 2-27.

Koomanoff, V.A. 1989. Analysis of global vegetation patterns: a comparison between remotely sensed data and a conventional map. Biogeography Research Series, Report #890201, Department of Geography, University of Maryland.

Korner, C. 1993. Leaf diffusive conductances in the major vegetation types of the world. in: *Ecophysiology of Photosynthesis*, E. D. Schulze and M. M. Caldwell (eds.) Ecological Studies 100, Springer-Verlag.

Kuchler, A.F.W. 1983. World map of natural vegetation. In: *Goode's World Atlas*, 16 ed., Chicago, Rand McNally, pp. 16-17.

Lambin, E.F. 1995. Change detection at multiple temporal scales: Seasonal and annual variations in landscape heterogeneity. *Photogrammetric Engineering and Remote Sensing*: in press.

Lambin, E. F. and D. Ehrlich. 1996a. The surface temperature-vegetation index space for land cover and land-cover change analysis. *International Journal of Remote Sensing* 17(3): 463-487.

Lambin, E.F. and D. Ehrlich. 1996b. Broad scale land-cover classification and inter-annual climatic variability. *International Journal of Remote Sensing* 17(5): 845-862.

Lambin, E.F. and D. Ehrlich. 1996c. Land-cover changes in Africa, 1982-1991. *Ambio*: submitted.

Lambin, E.F. and A.H. Strahler. 1994a. Indicators of land-cover change for change-vector analysis in multitemporal space at coarse spatial scales. *International Journal of Remote Sensing* 15:2099-2119.

Lambin, E.F. and A.H. Strahler. 1994b. Change-vector analysis: A tool to detect and categorize land-cover change processes using high temporal-resolution satellite data. *Remote Sensing of Environment* 8: 231-244.

Lee, J., W.C. Weger, S.K. Sengupta and R.M. Welch. 1990. A neural network approach to cloud classification. *IEEE Trans. Geoscience and Remote Sensing* 28: 846-855.

Lenney, M.P., C.E. Woodcock and H. Hamdi. 1996. The Status of Agricultural Lands in Egypt: The Use of Multitemporal NDVI Features Derived from Landsat TM. *Remote Sensing of Environment* 56 (1): 8-.

Lloyd, D. 1990. A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *International Journal of Remote Sensing* 11:2269-2279.

Loveland, T.R., J.W. Merchant, J.F. Brown, D.O. Ohlen, B.C. Reed, P. Olson and J. Hutchinson. 1995. Seasonal land-cover of the United States. *Annals of the Association of American Geographers* 85(2): 339-355.

Loveland, T.R., J. W. Merchant, D.O. Ohlen and J.F. Brown. 1991. Development of a land-cover characteristics database for the conterminous U.S. *Photogrammetric Engineering and Remote Sensing* 57: 1453-1463.

Ma, Z. and R.L. Redmond. 1995: Tau coefficients for accuracy assessment sensing data. *Photogrammetric Engineering and Remote Sensing* 61(4): 435-439.

Malila, W.A. 1980. Change vector analysis: An approach for detecting forest changes with Landsat, *Proc. 6th Ann. Symp. on Machine Process. Remotely Sensed Data, Laboratory for the Application of Remote Sensing,* Purdue Univ., pp. 326-335.

Malingreau, J.-P. 1986. Global vegetation dynamics: satellite observation over Asia. *International Journal of Remote Sensing* 7: 1121-1146.

Malingreau, J.-P. and C.J. Tucker. 1988. Large-scale deforestation in the Southeastern Amazon of Brazil. *Ambio* 17: 49-55.

Malingreau, J.-P., C. J. Tucker and N. Laporte. 1989. AVHRR for monitoring global tropical deforestation. *International Journal of Remote Sensing* 10: 855-867.

Malingreau, J.-P. and A.S. Belward. 1994. Recent activities in the European Community for the creation and analysis of global AVHRR data sets. *International Journal of Remote Sensing* 15(17): 3397-3416.

Matthews, E. 1983. Global vegetation and land use: New high resolution data bases for climate studies. *J. of Appl. Meteor.* 22: 474-487.

Mayaux, P. and E.F. Lambin. 1995. Estimation of tropical forest area from coarse spatial resolution data: A two-step correction function for proportional errors due to spatial aggregation. *Remote Sensing of Environment* 53: 1-16.

Mayaux, P. and E.F. Lambin. 1996. Tropical forest area measured from global land-cover classifications: Inverse calibration models based on spatial textures. *Remote Sensing of Environment*: in press.

McClellan, G.E., R.N. Dewitt, T.H. Hemmer, L.N. Mattheson and G.O. Moe. 1989. Multispectral image processing with a three-layer backpropagation network. *Proc. Int. Joint Conf. on Neural Networks* 1: 151-153.

Meeson, B.W., F.E. Corprew, J.M.P. McManus, D.M. Myers, J.W. Closs, K.-J. Sun, D.J. Sunday and P.J. Sellers. 1995. *Global Data Sets for Land-Atmosphere Models,* Volumes 1-5, Internat. Satellite Land Surface Climatol. Project, Initiative I, NASA Goddard DAAC Science Data Series, 5 CD-ROM disks.

Michaelsen, J., D.S. Schimel, M.A. Friedl, F.W. Davis and R.C. Dubayah. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. *Journal of Vegetation Science* 5: 673-686.

Michalek, J.L., T.W. Wagner, J.J. Luczkovich and R.W. Stoffle. 1993. Multispectral change vector analysis for monitoring coastal marine environments. *Photogrammetric Engineering and Remote Sensing* 59: 381-384.

Moody, A. and A.H. Strahler. 1994. Characteristics of composited AVHRR data and problems in their classification. *International Journal of Remote Sensing* 15(17): 3473-3491.

Moody, A. and C. E. Woodcock. 1994. Scale-dependent errors in the estimation of land-cover proportions: Implications for global land-cover datasets. *Photogrammetric Engineering and Remote Sensing* 60: 585-594.

Moody, A. and C. E. Woodcock. 1995. The influence of scale and spatial characteristics of landscapes on land-cover mapping using remote sensing. *Landscape Ecology* 10: 363-379.

Moody, A., S. Gopal, and A.H. Strahler.1996. Artificial neural network response to mixed pixels in coarse-resolution satellite data. *Remote Sensing of Environment* 58, in press.

Muchoney. D.M., J. Borak and A. Strahler. 1996. Global Landcover Classification Training and Validation Issues and Requirements. *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium*, Lincoln, Nebraska, May 27-31, Vol. 1, pp. 233-235.

Muchoney, D.M. and B. Haack. 1994. Change Detection for Monitoring Forest Defoliation. *Photogrammetric Engineering and Remote Sensing* 60(10): pp. 1243-1251.

Nelson, R. and B. Holben. 1986. Identifying deforestation in Brazil using multiresolution satellite data, *International Journal of Remote Sensing* 7:429-448.

Nemani, R. and S.W. Running. 1995. Global land cover changes and their impact on climate. *Clim. Change* 31: 395-413.

Nemani, R. and S.W. Running. 1996. Global vegetation cover changes from coarse resolution satellite data. *J. Geophys. Res*. 101: 7157-7162.

Nemani, R. and S.W. Running. 1996. Implementing a hierarchical global vegetation classification in ecosystem function models. *J. Veg. Sci.* 7: 337-346.

Nemani, R. and S.W. Running. 1996. Land cover characterization using multi-temporal red, nir and thermal-ir AVHRR data. *Ecol. Applications*. (In press)

Olson, J. S. and J. A. Watts. 1982. *Major World Ecosystem Complexes*. Oak Ridge National Lab, Oak Ridge, TN.

Pax Lenney, M. and C.E. Woodcock. 1996. The effect of spatial resolution on monitoring the status of agricultural lands. *Remote Sensing of Environment*: submitted.

Reed, B.D., J.F. Brown, D. VanderZee, T.R. Loveland, J.W. Merchant and D.O. Ohlen. 1994. Measuring phenological variability from satellite imagery. *J. Veg. Sci.* 5: 703-714.

Reich, P.B., M.B. Walters and D.S. Ellsworth. 1992. Leaf lifespan in relation to leaf, plant and stand characteristics among diverse ecosystems. *Ecol. Monographs* 62: 365-392.

Rosenfield, G.H. and K. Fitzpatrick-Lins. 1986: A coefficient of agreement as a measure of Thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing* 52(2): 223-227.

Roy, D. P. 1996. *The MODIS Land (MODLAND) Quality Assurance Plan*, Reworked draft version: September 20, 1996, 27 pp.

Rumelhart, D. E., G. E. Hinton and J. L. McClelland. 1986a. A general framework for parallel distributed processing. in: *Parallel Distributed Processing: Exploration in the Microstructure of Cognition. Vol. 1: Foundations,* J. L. Rumelhart, J. L. McClelland and PDP Research Group (eds.), MIT Press, Cambridge, MA.

Running, S. and J.C. Coughlan. 1988. A general model of forest ecosystem processes for regional applications: I. hydrologic balance, canopy gas exchange and primary production processes. *Ecological Modelling* 42: 125-154.

Running, S. and E.R. Hunt. 1993. Generalization of a forest ecosystem process model for other biomes, BIOME-BGC, and an application for global-scale models. In: *Scaling Physiological Processes: Leaf to Globe,* J. Ehleringer and C. Field (Eds.). Academic Press.

Running, S.W., C.O. Justice, V.V. Salomonson, D. Hall, J. Barker, Y.J. Kaufman, A. H. Strahler, A.R. Huete, J.-P. Muller, V. Vanderbilt, Z.M. Wan, P. Teillet and D. Carneggie. 1994b. Terrestrial remote sensing science and algorithms planned for EOS/MODIS. *International Journal of Remote Sensing* 15(17): 3587-3620.

Running, S.W., T.R. Loveland, L.L. Pierce and E. R. Hunt. 1994a. A vegetation classification logic based on remote sensing for use in global biogeochemistry models. *Ambio* 23: 77-81.

Running, S.W., T.R. Loveland, L.L. Pierce, R. Nemani and E.R. Hunt, Jr. 1995. A remote sensing based vegetation classification logic for global land cover analysis. *Remote Sensing of Environment* 51: 39-48.

Salu, Y. and J. Tilton. 1993. Classification of multispectral image data by the binary diamond neural network and by nonparametric, pixel-by-pixel methods. *IEEE Trans. Geoscience and Remote Sensing* 30: 606-618.

Sellers, P. 1991. Modeling and observing land-surface-atmosphere interactions on large scales. *Surv. in Geophys*. 12: 85-114.

Sellers, P.J., Y. Mintz, Y.C. Sud and A. Dalcher. 1986. A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci*. 43: 505-531.

Sellers, P.J., C.J. Tucker, G.J. Gollatz, S.O. Los, C.O. Justice, D.A. Dazlich and D.A. Randall. 1994. A Global 1o by 1o NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. *International Journal of Remote Sensing* 15(17): 3519 - 3545.

Settle, J. and N.A. Drake. 1993. Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing* 14(6): 1159-1177.

Shimabukuro, Y. . and Smith, J.A., 1991. The least-squares mixing models to generate fraction images derived from remote sensing multispectral data, *IEEE Transactions on Geoscience and Remote Sensing* 29(1): 16-20.

Smith, M.O., S.L. Ustin, J.B. Adams and A.R. Gillespie. 1990. Vegetation in deserts: I. A regional measure of abundance from multispectral images. *Remote Sensing of Environment* 31: 1-26.

Story, M. and R.G. Congalton. 1986. Accuracy assessment: A user's perspective. *Photogrammetric Engineering and Remote Sensing* 52: 397-399.

Stowe, D.A., R.L. Tinney, and J.E. Estes. 1980. Deriving land use/land cover change statistics from Landsat: A study of prime agricultural land. *Proc. 14th. Int. Symp. on Remote Sensing of Environment*, Environ. Res. Inst. of Mich., Ann Arbor, MI, pp. 1227-1237.

Strahler, A.H., A. Moody and E. Lambin. 1995. *MODIS Land Cover Product: Algorithm Theoretical Basis Document*, Version 3.0.

Townshend, J.R.G., V.Bell, A.C. Desch, C. Havlicek, C.O. Justice, W.E. Lawrence, D. Skole, W.W. Chomentowski, B.I. Moore, W. Salas and C.J. Tucker (1995). The NASA Landsat Pathfinder Humid Tropical Deforestation Project. *Land Satellite Information in the Next Decade*, Tyson's Corner, Virginia, American Society of Photogrammetry and Remote Sensing. pp. IV 76 -IV 87.

Townshend, J.R.G. and C.O. Justice. 1988. Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations. *International Journal of Remote Sensing* 9: 187-236.

Townshend, J.R.G. and C.O. Justice. 1990. The spatial variation of vegetation changes at very coarse scales. *International Journal of Remote Sensing* 11: 149-157.

Townshend, J.R.G., C.O. Justice, C. Gurney and J. McManus. 1992. The effect of misregistration on the detection of vegetation change. *IEEE Trans. Geoscience and Remote Sensing* 30 (5): 1054-1060.

Townshend, J.R. G., C.O. Justice and V. Kalb. 1987. Characterization and classification of South American land-cover types using satellite data. *International Journal of Remote Sensing* 8: 1189-1207.

Townshend, J., C. Justice, W. Li, C. Gurney and J. McManus. 1991. Global land cover classification by remote sensing: present capabilities and future possibilities. *Remote Sensing of Environment* 35: 243-255.

Tucker, C. J., B. N. Holben and T.E. Goff. 1984. Intensive forest clearing in Rondonia, Brazil, as detected by satellite remote sensing. *Remote Sensing of Environment* 15: 255-261.

Tucker, C. J., C.O. Justice and S.D. Prince. 1986. Monitoring the Grasslands of the Sahel 1984-1985. *International Journal of Remote Sensing* 7: 1571-1581.

Tucker, C.J., H.E. Dregne and W.W. Newcomb. 1991. Expansion and contraction of the Sahara Desert from 1980 to 1990. *Science* 253: 299-301.

Turner, M.G., R.V. O'Neill, R.H. Gardner and B.T. Milne. 1989. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology* 3: 153-162.

Tzeng, Y.C., K.S. Chen, W.-L. Kao and A.K. Fung. 1994. A dynamic learning neural network for remote sensing applications. *IEEE Trans. Geoscience and Remote Sensing* 32: 1096-1103.

Verbyla. D.L. and T.O. Hammond. 1995: Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified pixels with reference grids. *International Journal of Remote Sensing* 16(3): 581-587.

Walthall, C.L., J.M. Norman, J.M. Welles, G. Campbell and B.L. Blad. 1985. Simple equation to approximate the bidirectional reflectance from vegetation canopies and bare soil surfaces. *Applied Optics* 24: 383-387.

Wilson, M. and A. Henderson-Sellers. 1985. A global archive of land cover and soils data for use in general circulation climate models. *J. Climatol.* 5: 119-143.

Woodcock, C.E. and S. Ryherd. 1989. Generation of texture images using adaptive windows. *Tech. Pap. 55th Ann. Meeting Amer. Soc. Photogrammetry and Remote Sensing*, Baltimore, pp. 11-22.

Woodcock, C E. and A H. Strahler. 1987. The factor of scale in remote sensing. *Remote Sensing of Environment* 21: 311-332.

Woodwell, G.M., R.A. Houghton, T.A. Stone, R.F. Nelson and W. Kovalick. 1987. Deforestation in the tropics: New measurements in the Amazon Basin using Landsat and NOAA Advanced Very High Resolution Radiometer imagery. *J. Geophys. Res*. 92(D2): 2157-2163.

Yoshida, T. and S. Omatu. 1994. Neural network approach to land cover mapping. *IEEE Trans. Geoscience and Remote Sensing* 32: 1103-1110.

Zhu, Z. 1994. *Forest Density Mapping in the Lower 48 States: A Regression Procedure.* U.S.D.A. Forest Service, Southern Forest Experiment Station, New Orleans, Research Paper So-280, 9 pp.

7. Prior ATBD Reviews and Responses

7.1 EOS Review of June, 1994

The MODIS Land Cover/Land-Cover Change algorithm was reviewed by a panel of external referees in June, 1994, with the results made available in September, 1994. The ratings and comments of the panel are shown below. Panel members included Graeme L. Stephens, Pete Cornillon, Forrest Hall, Ben Herman, Hugh Kieffer, Teriyuki Nakajima, Jim Mueller, and Ichtiaque Rasool.

7.1.1 Ratings

Ratings of EOS Review Panel	
Criterion	***Rating****
1. Degree to which product meets EOS priorities	9
2. Soundness (feasibility/practicality) of approach	5
3 Appropriateness of algorithm input	5
4. Completeness of sensitivity and error budget	6
5. Soundness of validation strategy	6
6. Release of useful products at launch	9
**Rating scale: 9, high or strongly agree; 5, neutral; 1 low*	

7.1.2 Comments

By far the largest concern with this algorithm is the disparity between the simplicity of the classification technology used, and that available in the literature; for example, the NDVI-surface temperature decision tree classification method. A fixed threshold of 35 C is used to separate grass and croplands from wetlands and forest. Such a simple surface temperature threshold, even if properly corrected for atmospheric effects, cannot possibly work for all latitudes, at all times of the season. Surface temperature is quite variable and is basically not an intrinsic property of the surface. Therein lies its main weakness. The NDVI thresholds are ok as long as atmospherically corrected NDVI is used and background effects are accounted for.

Another concern is the ecosystem classes selected. They may be satisfactory as far as ecosystem process models go, but there are many other uses to which these classes will be subject. For example, finer partitioning may be needed for converting NDVI values to biophysical variables. Within the class of conifer defined by the Modis product, the relationship between NDVI and Fpar is quite variable. To accurately estimate Fpar for the conifer class for example, the class will have to be substratified into conifer subclasses for which the Fpar NDVI relationship is relatively constant. At this point, the Fpar can be aggregated up to any level desired by the ecosystem models. In addition to these concerns over the classes selected, other concerns were raised by reviewers that these classes might not be adequate for many energy, water, carbon cycling models.

No error budget was conducted here. To do so would require a system-level analysis in which errors in mapping ground cover areal extent and distribution are analyzed in terms of their effects on answering global questions. Quantifying classification at the pixel level does not provide much insight into performance because one does not know how such errors would propagate into regional and global scale questions. Another hindrance to validation is not knowing at regional and global levels just how good the information must be. What is missing in this regard is some accuracy criteria. One which might be useful is to assert that an algorithm is good enough when it produces significantly better information than that currently available. In the case of the land-cover product, global vegetation maps already exist. Will the MODIS-based algorithm produce better information? An issue is that it is not very clear how much better off the MODIS product will be over the circa 1978-mid 90's AVHRR product.

7.1.3 Response

The electronic mail message below, sent to Dr. Michael King, EOS Project Scientist, was prepared in response to the panel review of version 2.1 of the ATBD. It also includes some specific responses to comments of mail reviewers that are not reproduced here.

```
From alan@crsa.bu.edu Thu Dec 22 18:42:38 1994
Return-Path: <alan@crsa.bu.edu>
Received: from kalahari by crsa.bu.edu (8.6.4/Spike-2.1)
      id SAA20668; Thu, 22 Dec 1994 18:41:22 -0500
Received: by kalahari (931110.SGI/Spike-2.1)
      id AA01027; Thu, 22 Dec 94 18:41:21 -0500
From: "Alan Strahler" <alan@crsa.bu.edu>
Message-Id: <9412221841.ZM1025@kalahari.bu.edu>
Date: Thu, 22 Dec 1994 18:41:20 -0500
X-Mailer: Z-Mail (3.2.0 06sep94)
To: king@climate.gsfc.nasa.gov
Subject: Response to Reviews of MODIS Land Cover ATBD
Cc: fghall@imogen.gsfc.nasa.gov, lc@crsa.bu.edu,
        dhall@glacier.gsfc.nasa.gov,
        SWSHUETE@ccit.arizona.edu, justice@kratmos.gsfc.nasa.gov,
        jpmuller@ps.ucl.ac.uk, swr@ntsg.umt.edu,
        Vern_Vanderbilt@qmgate.arc.nasa.gov, wan@crseo.ucsb.edu
Mime-Version: 1.0
Content-Type: text/plain; charset=us-ascii
Status: O

Mike:

Version 3.1 of the MODIS Land Cover and Land-Cover Change ATBD is now
complete and deposited in various anonymous ftp directories as per your
instructions. I regret being unable to provide the revised document by
the November 1 deadline, but many, many hours of work were necessary to
revise the document both to respond to reviewers' comments and to accu-
rately reflect the current state of algorithm development.
In accordance with the instructions in your letter, this mail message
will address the changes in the document and provide specific responses
to the panel's and reviewers' comments.

Note that we plan a MODIS-IDS land cover workshop in late spring or
early fall to introduce IDS teams to the Land Cover Product. The objec-
tives are (1) to increase awareness of the product and its characteris-
```

tics among the IDS community; (2) to examine the suitability of the land cover classification units for IDS applications; and (3) to develop accuracy criteria and validation methodologies that are relevant to IDS science objectives. (The workshop might be combined with, or be part of, a broader workshop on MODIS Land Products, but this has not been discussed yet with MODLAND Team Members.)

OVERVIEW

In the period since the preparation of the original version of the document, the pathway to the land cover product has become much clearer. Major changes include:

1. Selection of a target system of land cover units for the Land Cover Parameter that follows the standards now being developed by the International Biosphere-Geosphere Programme (IGBP) for the IGBP-DIS (IGBP-Data and Information System) 1-km land cover database. These units are now being vetted by the IGBP Core Projects. Although this set of units may not respond to every need of every EOS investigator, it represents a solid starting point that is practical and will immediately satisfy a large portion of the scientific community. And, we will examine the nature of the units further at the MODIS Land Cover-IDS workshop mentioned above.

2. An emphasis on the use of statistical per-pixel classifiers in the land-cover classification process. This represents a step away from the physically-based NDVI-surface temperature thresholding algorithm that was presented at the ATBD review, and is a movement toward a statistically-based methodology that has the potential for higher per- pixel accuracy. Although the thresholding algorithm had the advantage of utilizing multitemporal NDVI and surface temperature data in a manner based on biophysical inference, the review panel felt that the procedure was too simple to provide good accuracy in a global application. Accordingly, we have refocused on statistical classifiers, especially neural network classifiers, for the Land Cover Parameter.

3. The processing path is specified much more clearly as a result of ongoing interaction with the MODIS Science Data Support Team and other Land Team members whose products will be used as inputs. This includes better specification of input and output data streams as well as improved estimates of processing needs and product sizes.

4. The post-launch product validation issue is now addressed more specifically in terms of the types of accuracy data that will be provided along with the product. The validation procedure will allow users to understand the uncertainty in the classification and model their own errors accordingly. The data will also provide subpixel cover type proportions that can be used in aggregation algorithms to provide accurate proportional covers at coarser resolutions than 1 km. Alternative validation procedures as proposed or conducted by IDS investigators will further be explored.

Before I respond to specific queries by the panel and individual reviewers, I want to make one general point: that Land Cover and Land- Cover

Change are empirical, data-dependent parameters. Thus, it is not possible to specify the exact algorithm to generate them without actually having the data in hand. In the prelaunch period, the task is to identify candidate algorithms as specifically as possible, given whatever test data are available that have MODIS characteristics. The objective is to be as far up on the curve as possible for a quick start with real MODIS data.

This is particularly true for the postlaunch Land-Cover Change Parameter. Since it will require at least two years of MODIS data to detect any sort of annual change, the product cannot be made until at least mid-2001. This is 6.5 years from now. During that time, we can expect that the state of the science will change significantly. At this point in time, our best strategy is to continue with the examination of key issues of change recognition at global scales using AVHRR rather than try to develop a specific MODIS algorithm.

REVIEW PANEL

Response to Specific Comments

1. Classification Methodology. The panel's “largest concern” was the use of the NDVI- surface temperature thresholding algorithm for land cover classification. As noted above in the overview, we have now moved away from this approach, and toward the use of neural network classifiers.

In defense of our earlier position, I should state that Steve Running and Rama Nemani, who are developing the thresholding approach to land cover classification, are working from a top-down perspective--that is, they are trying to get the broad patterns of land covers in their proper locations first, then worry later about per-pixel classification accuracies. This is certainly a valid approach, especially concerning the many problems with the quality of AVHRR 1-km composites at the pixel level.

However, further work by the BU team in the months following the ATBD review on datasets from small- (Plumas National Forest) and medium-sized areas (California) confirmed the panel's view that per-pixel accuracies of the thresholding method as currently devised are not sufficient for generation of a global product. Accordingly, we are now emphasizing the use of neural network classifiers for land cover. Because of our close working relationship with the BU Center for Cognitive and Neural Systems--one of the foremost research centers on neural networks and their applications--we are very well positioned to exploit the latest tools in this field for the MODIS application.

2. Land Cover Classes. Another concern was with the suite of land cover classes selected. For example, the review noted that it might be necessary to break up conifer (needleleaf evergreen forest) into subclasses in order to estimate FPAR accurately. (For this particular example, note that there is a MODIS product that provides FPAR directly- - MOD15.) The panel's concern has led us to two actions. First, we have provisionally adopted a set of land cover units for the product that is in the process of being vetted by the IGBP Core Projects for the IGBP- DIS global 1-km land cover database. Since the Core Projects include a very

broad cross section of researchers working in fields ranging from hydrology to carbon cycling, we think these units will be suitable for most IDS applications. Further, these units are designed to be recognizable with remotely sensed data, since they rely on structural and phenological distinctions that can be remotely sensed.

Second, we plan to conduct a workshop on the Land Cover Product (mentioned above) that will solicit further input from IDS groups on the exact nature of the classes and their needs for them. In this way, we will facilitate awareness of the product and enhance the feedback process by which we can best position our product to serve IDS needs.

Note that early in the process of developing the land cover product (1992), we conducted a survey of IDS PIs regarding their needs for land cover. The survey was not overly successful. At that time, few investigators had definite ideas about what their needs for land cover from MODIS were. However, among those who had some notion, nearly all endorsed our concept for a classification based on vegetation life-form that would include 20-30 global classes.

3. Validation and Errors. The panel noted the absence of an error budget, suggesting that to construct such a budget would require “a systems-level analysis in which errors in mapping ground cover areal extent and distribution are analyzed in terms of their effects on answering global questions.” They further remarked that knowing pixel- level classification accuracies “does not provide much insight into performance because one does not know how such errors would propagate into regional and global scale questions.” The panel also noted that accuracy criteria were missing, suggesting that one such criterion might be a comparison with accuracies of present land cover products.

A systems-level error analysis focused on answering global questions poses real difficulties for our effort. It would require using suites of models developed by other investigators and evaluating their sensitivity to error in land cover labels. We possess neither the expertise nor resources to carry out such analyses as part of the MODIS land cover activity. However, what we can do is to work more closely with IDS investigators and encourage them to assess the sensitivity of their own global studies to such errors. We plan to do so, beginning at the MODIS-IDS land cover workshop described above. The first step in this analysis, however, is understanding the types and frequencies of errors that may be anticipated--in short, the pixel-level classification accuracies mentioned above. This remains a key topic in our validation plan.

As for accuracy criteria, we may reasonably expect such criteria to come from IDS teams, especially after conducting sensitivity analyses as identified above. We will work with IDS teams toward that end. The panel's suggestion that the accuracy of present land cover products presents a practical standard for comparison is a useful one, and it can easily be implemented both during prelaunch algorithm development and postlaunch validation. Here, the specific target for comparison would be the IGBP-DIS global 1-km land cover database and the unsupervised clustering and labeling algorithm that will produce it.

Note that land cover in the presence of mixed pixels does not lend itself well to absolute accuracy standards, such as “90 percent correct.” When a pixel is 55 percent forest and 45 percent grassland, a label of forest may be technically correct, but misleading for many applications. This means that accuracy, labeling, and scaling processes all have to accommodate the mixed pixel problem.

Forrest Hall's E-Mail Comments

In the process of clarifying some of the Review Panel's comments, I contacted panel member Forrest Hall at GSFC. As part of our e-mail dialogue, he raised two issues about the Land Cover Product.

1. Spatial resolution. First, Forrest questioned the need for a product at 1-km spatial resolution, remarking that the finest resolution that global climate modelers have identified as usable within the next decade is 1/4 degree by 1/4 degree. I raised this issue with other members of the Land Team, and our consensus is that there are many regional applications of global land cover, ranging from carbon balance modeling to hydrology, that can easily exploit 1-km data. Perhaps the difference is that Forrest is looking at the problem from the ISLSCP perspective, while the Land Team is coming more from the IGBP perspective. The revised ATBD now contains a section on spatial resolution that cites the IGBP call for land cover at 1-km spatial resolution. It also points out that other 1-km resolution MODIS products will require a 1-km input.

Another important reason for a 1-km resolution product is that land cover proportions change with spatial aggregation. Smaller, more fragmented classes disappear with coarser scaling while larger, more abundant classes grow. If coarse-scale algorithms require land cover proportions as inputs, they will be wrong unless aggregation is carried out carefully to account for this effect. This requires initial classification at the finest practical scale, and for highest accuracy further requires subsampling at resolutions of about 100 m. These issues are discussed in more detail in the revised document.

2. Nature of Product. Second, Forrest raised the idea that the Land Cover Product would provide more than just land cover. He suggested it could be more like the ISLSCP Initiative 1 or 2 databases, which also include such parameters as monthly albedo, FPAR, LAI, biomass density, canopy green fraction, roughness length, standing water fraction, etc. Some of these parameters are inferred from land cover--for example, roughness length is typically inferred from vegetation height and cover. Others, such as FPAR, LAI, and canopy green fraction, are inferred from vegetation indexes, land cover, and a model (e. g., SiB2). Forrest generously sent page proofs of an article by Piers Sellers et al. in IJRS that provided an example of how such a database was derived at the 1-degree scale. Note that the MODIS Land Team provides a number of these as separately-derived products. Of course, this is quite a different concept for the Land Cover Product--one that is well beyond what was originally conceived or proposed.

My own position on this issue is that if ISLSCP-style modelers need a database of this nature, the Land Group should seriously consider pro-

viding it. With proper coordination among team members (and due attention to scaling issues), it would not be very difficult to aggregate our products into a monthly 1/4-degree by 1/4-degree (=772 sq km) grid. The products would include at least vegetation cover type, albedo, LAI, FPAR and VI, produced by the team. These products could then be merged with data that are derived from other models (e.g., SiB2) or are obtained by inference from the literature.
Note that Forrest and Piers anticipate a SWAMP/ISLSCP-sponsored meeting on the ISLSCP Initiative II Dataset in the late February time frame. MODIS will be represented. There will be also be a meeting of the Land Group, focusing on the interrelationships among land cover, BRDF/Albedo, and surface reflectance at BU in mid-January. I believe we can have a preliminary discussion of the issue at that time.

MAIL REVIEWS

Note that it is not possible to provide a document of a reasonable length and prepared in a reasonable amount of time that also addresses every possible concern or contingency that might be raised in the development of a global product such as land cover. The present revision of the ATBD focuses most directly on the primary issues raised by the Review Panel. However, there are a number of low-level issues that remain to be addressed more fully in later versions of the document and will either wax or wane in importance as product development proceeds. Some of these issues are cited by reviewers and discussed below. If not well defended in the document, at least they are discussed here and you may be sure that we are thinking about them.

Cihlar Review

1. Input Products and Compositing. Cihlar points out that the accuracy of the land cover product depends on the accuracy of the input products, and is concerned with the cloud screening and compositing process. Although we have addressed the compositing process in more detail in the revised ATBD, we are still trying to understand the cloud screening process better. Because the Land Cover Product will use surface reflectance as an input, the actual cloud screening will be done as part of that product. We are presently in a dialogue on this issue with the surface reflectance producers.

As for the accuracy of input products, other team members are still dealing with accuracy issues at this time, and it is not certain what error levels will be. However, we believe that our plan to exploit simultaneously spectral, spatial, temporal, and directional information is quite robust, and that as long as signal-to-noise ratios in input products are reasonable, our product will not suffer.

2. Land Cover Units. Cihlar suggests that MODIS' multiple spatial resolutions may impose a hierarchical classification structure; that the classification could be more detailed, with a larger number of units capable of being collapsed for different purposes; that mappings of the MODIS classification to others should be provided; and that the units need to be defined as objectively as possible.

As discussed above, a provisional set of land cover units, vetted by the IGBP, is now in place, and further examination of the classification scheme is ongoing.

3. Supervised vs. Unsupervised Classification. Cihlar notes that a combination of supervised and unsupervised classification techniques may be desirable to a reliance on supervised techniques alone. We believe that supervised techniques are better suited to our situation, in which we will have good information on land cover for training only at a limited number of sites. However, we remain open to using unsupervised techniques, as we state in the present version of the ATBD.

4. Land-Cover Change. Cihlar makes the points that the land cover change parameter should be sensitive to seasonal differences and be produced quarterly; may need to be combined with ancillary data in a GIS; and could provide unlabeled change vectors with translation tables, thus facilitating differing interpretation of changes. These are all good points that will be taken into account as the Land-Cover Change Parameter is developed. Note that the product is already planned for quarterly production.

Townshend Review

1. Classifier. Townshend suggests a two-stage classifier--a simple conventional technique first, followed by a more sophisticated classifier (e. g., neural net) on equivocal pixels. This is a good idea that is now incorporated in the document.

2. Input Data. Townshend raises several issues on the quality of input data--radiometric correction, geometric rectification, and atmospheric correction. We believe that the basic discussion of these issues belongs in the appropriate ATBDs for these algorithms and products. However, we are very cognizant that the quality of the Land Cover Product is dependent on the quality of its inputs. (See discussion under Cihlar 1. above.)

Loveland Review

1. Land Cover Variability. Loveland considers that the ATBD as reviewed pays too little attention to problems of spectral/temporal confusion, landscape complexity, and interannual variability of land covers. We agree that these issues need further development, and are not well addressed, even in the present version of the document. However, they are very data- and region-dependent. At this point, resources are not available for a global survey of the variability of specific land covers as a precursor to technique development. Our strategy is to address these issues as individual test sites within ecoregions come on line. Still, the caveat is an important one for allocating resources and anticipating problems.

2. Improvements in Product Quality with Time. Loveland suggests that although our first efforts may not be fully satisfactory, they will improve with time. This is perhaps a realistic assessment, especially considering the problems in land cover variability discussed above. How-

ever, we should still target for a high level of accuracy in the first products we prepare.

3. Land Cover Units. On this topic, Loveland notes that the list of land cover classes needs to be identified; land cover classes need to be related to the biophysical characteristics of the land cover; classes need to be linked with types in current use to provide continuity; and that classes need to include logic for handling mixed covers and heterogeneous landscapes. All these concerns are satisfied with the present product plan as described in the ATBD.

4. Ancillary Data Specifications. Loveland suggests that specifications for ancillary data (climate and DEM) need to be provided in more detail. We have dropped the need for a climate database, instead relying on an ecoregion-based processing scheme to make the broad distinctions that we hoped to gain from the climate database. As for the DEM, it will be used only as an elevation variable to distinguish broad land covers with similar signatures but different topographic habitats. For this purpose, any reasonable accuracy standard should suffice.

5. Land-Cover Change Parameter Classes. Loveland points out that we have focused primarily on issues of landscape biophysics for this parameter, while the application of the product is likely to be broader, including issues of human modification, adaptation and mitigation. Inasmuch as the focus of the EOS global change agenda is on global climate change and not on human activity, I believe that our focus is properly placed. However, human activity has undoubtedly altered our planet's climate and will continue to do so in the future. This means that our product must be sensitive to those anthropomorphic processes that influence climate. We plan to ensure that such a focus is maintained as the Land-Cover Change Parameter is developed further.

Field Comments

1. Neural Net not Mature. Field points out that the neural net classifier we proposed is not mature and could require further development. We concur, and development is ongoing.

2. Land Cover Units: Field expresses a number of concerns with land cover units. He notes that the units are not suitable for all classes of models for which they might be used and suggests that several different classifications might be carried out and presented as alternative products. These could range from a very simple set of units that might be suitable for climate models to a system of many units (hundreds) for use in modeling nonmethane hydrocarbon emissions or soil emissions of nitrous oxide. A further alternative is not to classify at all, but to produce an output of continuous variables associated with land cover type, such as woodiness, leaf longevity, or fraction of perennials.

These are interesting ideas that deserve consideration. It is certainly true that no set of units will suffice for all applications, and this issue will be aired more directly with IDS investigators. Carrying out multiple classifications adds a degree of cost and complication beyond that currently planned.

It is hard to envision a system of hundreds of units that are recognized on a global scale and that are also remotely sensible with good accuracy. That is, there is no point in setting up a system of units (e. g., a floristic plant classification) that does not have an inherent physical basis for a differential response in the emission and scattering of electromagnetic radiation, or in the temporal, spatial or directional distribution of that radiation, that is sensible with the instruments at hand. If such a procedure works, it will only be for a limited number of classes that possess attributes of spectral signature, structure and phenology that are coincidentally correlated with the insensible attribute. In our view, it is better to recognize a smaller set of more broadly-defined units with good accuracy than a larger set of finely-defined units with poor accuracy.

One possibility would be to follow the example of Loveland's classification of the conterminous U. S. from NDVI composites. In this work, a large number of classes are recognized by unsupervised clustering and classification, and then collateral data are used to characterize them. The problem with a global approach of this nature is that the interpretation is dependent on the quality of the collateral data, which for many areas will be nowhere near the standard for the U. S. Perhaps it would be possible to simply provide the class map without labels, and allow individual investigators to label them based on their own knowledge and the inferences they wish to make. In any event, these issues will be discussed directly with IDS investigators in the future.

Field's idea that continuous variables could be output instead of categorical classes is an interesting one, but it would require much development. To our knowledge, no algorithms exist for such products. If this idea captures the interest of IDS investigators, it could be pursued as a postlaunch activity.

Alan Strahler
Center for Remote Sensing
Boston University

7.2 SWAMP Review of May, 1996

The MODIS Land Cover/Land-Cover Change algorithm was reviewed by a panel of external referees in May, 1996 convened by the Science Working Group, AM Platform (SWAMP). Results were made available in early October, 1996. The ratings and comments of the panel are shown below. Panel members included Marvin Bauer, Josef Cihlar, Robert Davis, Narendra Goel, Yan Kerr, John Miller (Chair), John Mustard, and John Price.

7.2.1 Ratings

Ratings of SWAMP Review Panel			
	Rating[1]		
Criterion	*Land Cover*	*Land-Cover Change*	*Townshend Proposal*
1a. Technical/scientific soundness of algorithm/approach	7/3[2]	8	8
1b. Value of the data product to the land science community	2	7	9
1c. Soundness of the validation strategy	7/2[2]	6	7
1d. Extent to which 1994 review issues have been addressed	8	N/A	N/A
2a. Extent to which product is (1) useful to broader needs of community and (2) meshes with other data products	7	6	7/0[3]
2b. Assessment of plans for comparison or enhancement of similar data products from other instruments	4	4	2

[1]*Rating scale: 9, excellent, strongly agree, high; 5, average, neutral, medium; 1 poor/needs work,disagree, low; 0, insufficient information.*
[2]*Second value is for land cover algorithm used by Running in LAI/FPR product.*
[3]*Values refer to two parts of criterion.*

7.2.2 Comments

The text of the comments of the panel is reproduced below. (Note that section numbers below are those of the SWAMP report, not this document.)

7.2.5a. Data Product: MOD12 - Land cover type

Strahler et al.

(Review based on ATBD-MOD-13 dated November, 1994
and presentation at workshop May 16, 1996)

(Comments are also made on the work of Running et al:
ATBD MOD17
Presentation 05/96)

(a) technical/scientific soundness of the algorithm/approach described (Rating: 7/3)

The basic approach consists of: rigorous processing of MODIS data (from other MODIS products), using of a variety of data sets to provide as much ancillary information as possible to improve chances of correct classification, using new recent classification algorithms trained on known sites, and validation at various sites around the world. In principle, the approach is sound. However, a number of issues need to be dealt with:

(1) The neural network classifiers do not permit, or at least they make difficult, insights into the classification process and reasons for misclassifications. Their performance is limited by the accuracy and representativeness of the training data which are likely to be deficient in many parts of the world (see below). They are likely to perform well over areas similar to those used to train and are more likely to “blow up” outside of these areas. Note that the reference for ARTMAP is from a simple land environment (Sahel, p.19). Hara et al (p.19) had mixed results with neural net classifier, and had to add Maximum Likelihood (ML) classifiers to make it work for only three, quite distinct cover types. The accuracy results presented in 05/96 (for this and other classifiers) are not particularly encouraging.

Use of the at-launch product of EROS/IGBP is acceptable but land cover data from a range of test sites should be used to evaluate and document its accuracy.

The proposed resolution of 1 km negates a major advantage offered by MODIS over previous data sets - higher spatial resolution. Even with AVHRR 1 km has consistently been the goal because land cover is highly heterogeneous at the sub-km scale and so are the processes (especially where anthropogenic effects are significant). The arguments raised on p. 24 are true but there are other sources of error. Strong effort should therefore be made at producing a higher resolution product (at least 500m).

The proposed data set (3.1.1.1.1) consists of products representing various time periods and potentially changing surface conditions. For example, surface reflectances will be obtained once per month (from any date, thus allowing a wide range of reflectance for similar cover types, e.g. in cropland, northern areas,.. - the compositing procedure envisioned is not spelled out, and has not been discussed in the published literature - it seems that the team proposes to do the compositing, p. 25); VI also once per month (for the same data? not clear; will the VI be computed from the surface reflectances?); snow/ice cover from weekly maps could represent a different date than surface reflectances or NDVI; same for surface temperature. In using such a potentially disparate data set there is a danger of strong noise overwhelming (often weak) cover type signal. The data set is to have 18 variables per pixel. (From the presentation it was not completely clear which of the above are still considered for the product, and which have changed).

It is proposed to use data from 12 months. In areas with seasonal snow cover this is likely to present a significant problem because seasonal snow is not an inherent land cover characteristic, has a marked effect on the measured signal (as well as its BRDF properties etc.), and tends to be highly variable. The ATBD does not explain how this will be dealt with. One approach is to screen snow-covered parts of the data out (Cihlar et al. RSE 1996, in press).

Strong use of temporal information will increase the likelihood that the classification reflects the seasonal dynamics of the year when the data were obtained (regional droughts,...). An option is to de-emphasize the temporal dimension in the data set, e.g. through extraction of intermediate/higher level variables.

Regarding use of BRDF parameters for the classification: this seems reasonable but was it tested, and with what results?

The supposition that dividing the land surface into ecoregions for classification will increase efficiency needs verification. It has the disadvantage of having to resolve misclassifications at the edges, with attendant inconsistencies and potentially arbitrary rules. Ecoregion boundaries are often arbitrary and the reconciliation could lead to much additional work. It would seem preferable to have a consistent approach for as large areas as possible, at least a continent.

It is proposed to train the statistical classifiers with data from 40-60 sites around the world. This approach has some pitfalls: it limits the accuracy of the final product because such classifications (especially if derived from satellite data such as LANDSAT) are only 70-90% accurate; it assumes that equivalent “map” (same spatial resolution and cover types) can be derived from the high resolution data to train the MODIS data - this is not straightforward; it assumes that the detailed classification is valid for the period of MODIS data coverage - this may be difficult to achieve everywhere; it requires a sufficiently large sample of “pure”, unmixed pixels for the various cover types (especially as the proposed number of input channels is large) which may be difficult to achieve in many situations; and it assumes that signature extension will work, and this has generally been difficult to achieve in most studies to date.

As an alternative to the complex classification approach the team should consider the use of a “small” number of intermediate variables which describe the surface characteristics (e.g., Defries et al. 1995, RSE 54: 209), understand the land cover dependence of these variables, and use relatively simple classifiers (not forgetting about minimum distance) to produce classification results. Note some of the most detailed results so far have been achieved with single date images (AVHRR channel 1 and 2; Beaubien and Simard, 1993, Methodologie de classification des donnees AVHRR pour la surveillance du couvert vegetal. Proceedings of the 16th Canadian Remote Sensing Symposium, Sherbrooke, Quebec: 597-603.)

Re: Classification by Running et al: (Rating: 3)

It is proposed to use a simple classification scheme, designed to respond to the needs of BGC modelers. The classes are intended to be extractable from remote sensing data. The classification approach features a small number of classes and fixed thresholds to separate these on a global basis.

The use of a fixed set of thresholds globally to separate these classes with a reasonable accuracy is not likely to succeed because of the variety of environments. A comparison with the USGS land cover map quoted at the presentation (70%) bears this out (although it should be noted that accuracy of the latter is not known precisely either). This is particularly true for surface temperature. In defining a more specific set of thresholds for different regions/biomes one must get closer to the approach followed by others (e.g., Strahler et al., ..) which raises the possibility of several different products.

(b) value of the data product to the Land science community (Rating: 2)

Land cover is a critical parameter for most land applications. Its value would be significantly enhanced if produced at a higher spatial resolution. If 1 km is adhered to one can only hope for a better accuracy compared to pre-launch products.

(c) soundness of the validation strategy (Rating: 7/2)

Re Strahler et al: (Rating: 7)

Sources of error/uncertainty: good discussion. The cloud contamination problem depends in part on the quality of cloud screening (will be especially challenging for land due to broken clouds, snow, haze/smoke) and on the ability to compensate for missing data. The ATBD does not discuss how to deal with missing data. NN classifiers may be able to do that but there should be a backup plan. Interpolation is a possible option (Cihlar et al., RSE 1996 in press).

Another factor not considered in sufficient detail is surface water class. The proposed input data will not provide useful information about water, and another approach is required. A water mask would be ideal but unless it is availability at a sufficient resolution worldwide a parallel methodology (based e.g. on a one-time, single-date, cloud-screened images) should be examined.

It is very unlikely that 100% accurate reference data will be available to the team. Thus the specified product accuracy will be obscured by this uncertainty. This needs to be worked out and the implications for the users ascertained.

The concern about mountainous areas is well placed but not enough information is provided about the proposed solutions.

The idea of fuzzy labeling has some merit but it should be noted that users will need definitive labels (e.g., this pixel has 45% of A and 55% of B), which comes back to accurate classifier in the first place. The problem will be mitigated if smaller pixels are used.

Re Running et al: (Rating: 2)

No detailed validation strategy has been described. Broad comparisons with other global data sets (whose accuracy may not be well known) is not sufficient.

(d) extent to which 1994 ATBD review issues have been addressed (Rating: 8)

The approach of Strahler et al addressed many of the issues raised by the ATBD review. The adequacy of the proposed classes in terms of the needs of carbon, energy, and water cycle modelers has not been assessed, nor has the classification accuracy required by these models. Recognizing limitations in the ability of satellite data to provide comprehensive information for modeling purposes (e.g., wetlands) the implications need to be examined. This requires collaboration between the classification teams and various modeling groups.

(e) near-term recommendations for improvements to the data product

- It is recommended that one MODIS-based land cover product be generated as soon as possible after launch, as detailed (including pixel size) and accurate as feasible given the available data and knowledge of information extraction techniques.

- The product should have categories that can be characterized and/or generalized for use in the various carbon, energy, water cycle models. This product should be comprehensively validated with 100% accurate data where possible. For pre-launch, the product based on 1992 global 1 km AVHRR data can be employed. However, effort should be made to define this product in more detail and proceed with a pre-launch validation program.
- MODIS scientists should make efforts to collaborate with IDS teams and others outside EOS in comparing classification strategies and validating candidate products in a range of environments.

(f) long-term recommendations for improvements or additions to the data product

- The classification product team should collaborate with the modelers to determine the impact of the classification on the modeling results, and to plan future improvements in the classification scheme. Other user communities should be canvassed for their needs.
- The highest spatial and thematic resolution feasible at the global level should be aimed at. In developing future products other related activities, both within EOS and by other groups, should be taken into consideration.

7.2.5b. Balance of Land Data Products as generated by EOS-AM 1 (i.e. ASTER, MISR, MODIS) to meet the needs of the broader Land science community

(a) extent to which the data product (and its accuracies) is useful to the broader land science community and meshes with the other instrument data products (Rating: 7)

This will be a very useful product, provided it is spatially detailed, has adequate land cover categories (thematically rich, perhaps 15 to 20 initially), and is accurate.

(b) assessment of plans for the comparison or enhancement of similar data products from the other instruments? (Rating: 4)

This product will be unique at the global level.

(c) recommendations for changes to improve the balance of land data products

See recommendation above (one pre-launch, one immediately post-launch product).

7.2.6a. Data Product: MOD12a Land cover change

(Review based on ATBD-MOD-13 November, 1994 and presentation at workshop May 16, 1996)

(a) Technical soundness of the algorithms/approach described: (Rating 8)

The 1994 ATBD noted the preliminary nature of the description provided. Five approaches were described at the presentation but the algorithm has not yet been finalized. The various approaches are all likely to produce information on change but they have different strengths and weaknesses, and would probably be optimum for different environments. For an EOS land cover change product the consistency, stability and "meaning" (information content) of the product should all be relatively stable.

Note that the derivation of this product should be based on fully-corrected MODIS data at the highest resolution feasible (preferably 250m).

(b) Value of the product to the Land science community: (Rating: 7)

This product will give a general idea of changes occurring globally over land in a given year. It should therefore be useful for public policy discussion and general awareness of the changes occurring. This assumes that the changes can be correctly classified, not just detected and quantified. From the scientific perspective the value of the product is less clear. It should be useful for stratification purposes, as a tool to focus further efforts etc., but further exploration of the changes indicated by this product must be based on other (lower level or other types of) products and models. This product may serve for preliminary/independent validation of the models outputs.

(c) Soundness of the validation strategy: (Rating: 6)

It is not clear what is validated. The chosen algorithm will produce a result, and there is no equivalent surface variable to compare with. Only the type of change can be validated once it is detected. Thus it is not guaranteed that the changes will occur at the LANDSAT Pathfinder sites, although some are bound to take place there, given 120 locations. Same comment for the IGBP sites. This is in addition to the (anticipated by ATBD) inferior quality of some of the surface information. Especially important in this respect is up-to-date satellite data coverage and recent (post-launch) surface data.

The post-launch validation strategy (reliance on more detailed satellite and other) is reasonable but the plans are not well developed at this stage.

(d) Extent to which 1994 ATBD issues have been addressed: (Rating N/A)

No issues raised in 1994.

(e) Near-term recommendations for improvements to the data product

- Define one (or a limited number) of algorithms to be used, and thoroughly evaluate their characteristics/peculiarities so that useful products can be generated shortly after launch.

(f) Long-term recommendations for improvements or additions to the data product

None.

7.2.6b. Balance of Land Data Products as generated by EOS-AM 1 (i.e. ASTER, MISR, MODIS) to meet the needs of the broader Land science community

(a) extent to which the data product (and its accuracies) is useful to the broader land science community and meshes with the other instrument data products (Rating: 6)

This product should be very useful to the earth science community, for stratification, planning field experiments, and assessment of the dynamics of the global ecosystems.

(b) Assessment of plans for the comparison or enhancement of similar data products from the other instruments (Rating: 4)

This is a unique product. Data from other instruments (ASTER, ETM) will be used to analyze the detected changes.

(c) Recommendations for changes to improve the balance of land data products

None.

7.2.7a. Data Product: MODLANDCOVER - (New proposal by Townshend)

(Review based on presentation and summary notes at workshop May 16, 1996)

General comments:

This product is brand new and complements the ATBD land cover and land cover change. It is based on a much more classical approach and on a good experience of land cover classification, though somewhat restricted to a few ecosystems. The aim here is to prepare an at-launch research product, and to provide help for the land cover ATBD. We can only encourage this, considering what we have seen of the planned product.

Also a definition of what is meant exactly by land cover change would be useful; if it is obvious for deforestation it is much less so for desertification or land use changes.

(a) technical/scientific soundness of the algorithm/approach described (Rating: 8)

It must be said at the start that the proposed algorithm is more research-oriented than “operational”. Consequently there are several issues to be dealt with and the author seems to be well aware of the different caveats. The main advantage of the proposed approach is (a) to effectively use the 250 m resolution, (b) to investigate other classifiers which are less exotic and already tested (Markov's for instance). The authors will rely on other MODIS products (cloud mask for instance) but with care which seems to indicate that he had already evaluated the product. The points of a little concern are linked to the use of thresholds (depending on viewing conditions, environment, errors in the processing chain,....) and the choice of and taking values within a window. Subtle changes in vegetation cover do occur over small gradients which might give way to a blurred response over too large a window, making it undetectable. However, the authors seem to be aware of the problem, as they are aware of the land cover change detection issue of mixing seasonal an inter-annual variations with actual land cover change. A good ratio of past expertise and test of new methods with some pragmatism is recommended. Mixture modeling, Markov's random fields and image segmentation are quite promising from the experience of others in the field. In addition, the temporal aspects could be used more, as they have proven to be extremely useful in both phenology monitoring (and hence vegetation type identification) and effective cloud decontamination.

(b) value of the data product to the Land science community (Rating: 9)

There is no question on the usefulness of such a product for the land science (and application!) community. Land cover is a crucial product for both other algorithms and for most land application (NPP, carbon and water cycle etc...). As the authors propose to deliver a product at 250 m it will be all the more useful.

(c) soundness of the validation strategy (Rating: 7)

The validation strategy is sound but seems to be very ambitious. One might have expected to see more IDS and eventually other instrument team members involved. The list of persons given in section 12 corresponds to busy individuals who might not be able to devote a sufficient amount of time to the validation process. Anyway for such a product, validation is not straightforward and the use of LANDSAT to assess changes is probably the way to go. Validation will nevertheless remain an issue for all the remote areas. The meaning of accuracy, even at 250 m is not clear.

(d) extent to which 1994 ATBD review issues have been addressed

Not applicable.

(e) near-term recommendations for improvements to the data product

- The data product is not yet fully mature as the proposal has just been put out. Except for what was written above, the main issues seems to have been addressed and we encourage the team to follow the steps indicated in their proposal.

(f) long-term recommendations for improvements or additions to the data product

- Not really applicable here as the intent is, we believe, to eventually merge the efforts with the current land cover/change product.

7.2.7b. Balance of Land Data Products as generated by EOS-AM 1 (i.e. ASTER, MISR, MODIS) to meet the needs of the broader Land science community

(a) extent to which the data product (and its accuracies) is useful to the broader land science community and meshes with the other instrument data products (Rating: 7/0)

The product will be useful for the community at large, at least for all studies/instruments with medium to low resolution, as long as it will be “accurate enough” which will probably be the case. (Rating: 7)

However we notice that there is almost no reference to other instruments (MISR, ASTER) as inputs and that there is no discussion on how to intercompare/standardize products. (Rating: 0)

Otherwise, the proposers have close relationships with international programs which is certainly a plus.

(b) assessment of plans for the comparison or enhancement of similar data products from the other instruments? (Rating: 2)

Apart from LANDSAT (and this point is quite interesting and promising) there are NO plans to interact with other instruments.

(c) recommendations for changes to improve the balance of land data products

We would strongly suggest to the proposers to:

- Standardize inputs/outputs with other instrument teams
- Develop a common strategy for validation sharing the efforts and costs (instrument teams and IDS)

7.2.3 Response

The following text was prepared in response to the specific comments of the review panel, which are reproduced above.

I. Technical/scientific soundness

1. The panel notes several limitations of neural network classifiers, including (1) difficulty in deriving insights into the classification process and reasons for misclassifications and (2) dependence on the quality of the training data. They also note that (3) the test of ARTMAP cited in the ATBD is overly simple and (4) the accuracies presented at the review on 5/96 are not particularly encouraging.

Every classification method has advantages and disadvantages. The main strength of neural network classifiers is that they provide a disjunct, curvilinear partitioning of measurement space that can follow training data more accurately than conventional multivariate partitionings, such as those of the maximum likelihood classifier or discriminant functions. Generally, this behavior will be an advantage, rather than a disadvantage, for achieving high classification accuracy. Another advantage is that, once trained, neural network classifiers run much more quickly than conventional multivariate classifiers, which is an important consideration for a global processing algorithm.

The internal structure of a trained neural network contains the information used to discriminate among classes, but since it is carried in the weights within the neural feedback structure, it is typically difficult to analyze and understand. However, some progress is being made. Gopal and Woodcock (1996) analyzed the internal structure of a neural network using principal components of network weights in a study involving change detection. The study reveals that the neural network exploits the same identifiable scene characteristics as other multivariate change detectors, such as Gramm-Schmidt orthogonal functions. Note also that this area is receiving a lot of attention within the neural net community, and we can expect good progress to continue in the near future.

As is the case for all classifiers, the representativeness of the training data is an important factor in achieving high accuracies. Yet this is not the overriding criterion, as many studies have shown that classification accuracies increase when neural net classifiers are simply swapped with a maximum likelihood classifier trained using the same data. Jackknifed trials of varying training/testing data subsets show consistent results, indicating that accuracy gains are not overly sensitive to the exact selection of training data. Moreover, there are approaches that can help overcome the limitations of specific training sets, such as that of Brodley and Friedl (1996), described more fully in the ATBD, in which training pixels that are classified equivocally by multiple classifiers are discarded as outliers.

In the period since the preparation of the prior ATBD, we have expanded our testing of the neural net classifiers, and confirmed their superior performance with respect to maximum likelihood. These studies include the classification of land covers (a) in the Sahel using 1-km AVHRR data (previously reported); (b) in Austria using TM data (Fischer et al. 1996); (c) in Walnut Gulch/Cochise County, Arizona; and (d) globally using composited AVHRR NDVI data at 1-degree spatial resolution. In each case, the neural network classifier significantly outperformed the maximum likelihood classifier used as a control by 5 to 15 percent. We have also included studies of decision tree classifiers in cases (c) and (d) and shown them to be on a par with neural nets, thus providing an alternative approach that may ultimately outperform neural net learning.

We are surprised at the review panel's remark that the results presented in 5/96 are not particularly encouraging. The classification accuracies using neural networks and decision tree classifiers that we are typically in the 85 to 90 percent range (see section 3.1.4). Given the diversity of signatures of the land cover types, we believe that these results are very good, and represent the state of the art for classification of coarse land covers using satellite data. Our presentation did cite some results for the Walnut Gulch/Cochise County study showing accuracies around 75 percent; however, we believe these are low due to overly generalized labeling of the training and testing data (see section 3.1.4.1.1).

Overall, our strategy is to develop neural network and decision tree approaches as far as practical in the prelaunch era so that we are prepared to use the best, most effective “advanced technology” classifiers as MODIS data become available. Note that it would be easy to step down to more conventional techniques in the postlaunch era, should they be preferred, whereas it would be very difficult to try to adopt new approaches without specific experience with them.

2. The panel states that the use of the EROS/IGBP dataset is acceptable, but that its accuracy should be evaluated and documented.

Agreed. As part of his effort to provide the at-launch dataset, Townshend is developing a plan to assess accuracy of the at-launch product. The accuracy of the IGBP product is the subject of a careful study proposed by Jack Estes of U. C. Santa Barbara, with Strahler as a Co-Investigator. If funded, it will provide the best documentation of accuracy for a global product ever attempted. It will also provide the opportunity to acquire a validation database that will be of immediate application to the MODIS product.

3. The panel notes that MODIS provides a new source of global data at higher spatial resolution and requests that strong efforts be made at producing a product at 500 m or finer.

There are a two major technical obstacles to providing data at finer resolution than 1 km. First is the problem of geolocation accuracy and change in pixel size with scan angle. As documented in the ATBD, expected geolocation error is +-85 m along track and +- 153 m across track at nadir (3 sigma values), and pixel size increases to 1103 m by 2415 m at ±55° scan angle. These values provide an effective IFOV of the instrument's 500 m bands of 656 m along track and 960 m across track when compositing samples over time. Moreover, no single date image will overlay that of a prior date exactly, so interpolation (and consequent smoothing) is required. Thus, the information content of the data does not support 500 m resolution, especially in the across-track (along-scan) direction.

Second is the problem of computation. A halving of spatial resolution produces a fourfold increase in processing time and a fourfold increase in the size of intermediate datasets. The issue is not so much in the classification, which is done infrequently, but in the resolution of the input data. The land cover algorithm uses spectral data that are corrected for atmospheric and view angle effects by the BRDF/Albedo algorithm. To produce these data at finer resolution means running this code at finer resolution, which quadruples the processing time and data volume. This algorithm is already one of the “tall poles” in MODIS Level-3 processing, and it will be some years after launch before EOSDIS will be able to accommodate such a computation load. (Note that current estimates are that EOSDIS will not be able to provide products much beyond Level-1 at launch, and that it will be some time before global Level-3 products are made.)

These factors will prevent us from making a finer resolution product on the global scale in the immediate postlaunch era. However, it would be possible to make a 500-m research product for some selected regions as a way of developing appropriate algorithms for a next-generation product. This would entail individual rectification of near-nadir image swaths from clear dates to build a suitable high resolution database. We will investigate this possibility and try to provide such products to test interest and demand.

4. The panel observes that a number of details of the compositing procedure are not specified in the ATBD they reviewed. Moreover, they are concerned that data for a grid cell could come from any date within the

32-day compositing period, thus introducing temporal noise into the data.

The compositing process is now more fully documented in the present version of the ATBD.

When multiple values of an input parameter are available, the compositing process will select the one with the best quality, and thus there may be temporal noise in the dataset. However, a key attribute of both neural net and decision tree classifiers is that they are robust and handle noise gracefully. In fact, this is one of the reasons for selecting these algorithms for the classification procedure. In addition, both neural nets and decision trees perform feature selection as part of the training process. In the case of decision trees, the selection is explicit and based on a number of rather sophisticated algorithms; for neural nets, the selection is implicit by weighting some features more heavily than others. It is possible that some filtering of input data will be useful and needed. In that case, we can look to such preprocessing procedures as FASIR (Los et al., 1994) as prototypes. Note also that considerable noise reduction will occur in the view angle correction for BRDF/Albedo, which fits a semiempirical model to the angular distribution of observed surface reflectances.

5. Seasonal snow cover can present problems, according to the panel report, since it is not an inherent land cover characteristic and will be highly variable.

To some extent, variation in snow cover within land cover type is temporal noise, and thus the considerations of the previous point apply. However, for the classifier to see the snow cover as noise, training sites need to reflect the full range of snow cover dynamics within types. We will endeavor to select such a range of sites. Snow-screening, suggested in the report, is also a possibility.

I should also note that we are doing algorithm development and testing with the Snow/Ice Team using data from Glacier National Park, and that we have a joint field campaign at the Hubbard Brook/Sleepers River test area in Vermont/New Hampshire scheduled for February. It will include MAS overpasses.

6. The panel expresses the concern that strong use of temporal information will increase the likelihood that the classification reflects the seasonal dynamics of the year in which the data were obtained, citing the case of seasonal droughts. They suggest that intermediate/higher level variables may reduce the dependence on temporal information.

We don't believe this will present a problem, since the information will be in both the training and classification data. As in the preceding point, representative training data are important.

Intermediate and higher level variables, such as maximum, minimum, range or time integral of NDVI in an annual cycle, have been used in some classification studies (Lloyd, 1990; Moody and Strahler, 1994; DeFries et al., 1995; Reed *et al.*, 1995), primarily to reduce data volume. They can easily be used in our algorithm if they prove as effective as original data. However, it is not obvious what variables to use for such inputs as surface reflectance. Note also that feature selection using decision trees (see section 3.1.4.1.1) can also be effective.

7. The panel wonders what tests we have carried out on the use of BRDF parameters in classification and what were the results.

This is an area that we are continuing to work. The exact parameters to be used are still TBD, but we expect them to convey information about the vegetation structure either through the specific surface scattering model selected and/or by the weights of the parameters fitted to the scattering model. We do know, however, that important information is present in the directional signal. Section 2.2.3 cites several recent studies that demonstrate the principle. Also, at the Beijing BRDF workshop (September, 1996), Marc Leroy presented some impressive results for distinguishing forest types at BOREAS using multiangle data from the POLDER simulator.

8. The panel observes that the use of ecoregions for processing will most likely produce more problems than it solves.

On reflection, we agree, and will instead process by continents.

9. The panel points out some problems with the supervised classification approach, which requires training sites. They note that this approach (a) “limits the accuracy of the final product because such classifications ... are only 70-90% accurate;” (b) “assumes that ... [the] same spatial resolution and cover types can be derived from the high resolution data to train the MODIS data - this is not straightforward;” (c) “requires a sufficiently large sample of 'pure,' unmixed pixels for the various cover types ... which may be difficult to achieve in many situations;” and (d) “assumes that signature extension will work, and this has generally been difficult to achieve in most studies to date.”

As a general response, we have selected a supervised approach largely because it is the most cost effective. It is easy to work with a supervised classifier quickly and efficiently, iterating classification parameters and training data in a logical, straightforward fashion. Unsupervised classification, which is the other alternative, requires a long process of stratifying and labeling classes. We believe that to do land cover faster, better, cheaper, we will need to use the supervised approach, and also believe that advanced supervised classifiers are up to the task. These classifiers can partition measurement space in ways equivalent to the labeling of large numbers of unsupervised classes, thus blurring the distinction between supervised and unsupervised classification. We are also exploring some unsupervised classifiers, including a Gaussian version of ARTMAP.

With that said, however, it is certainly true that the volume and quality of training site data will be a critical factor in delivering an accurate product. We have revised our approach to training sites significantly since the writing of the ATBD that was reviewed. We now have a strategy of collecting many more training sites and maintaining several different levels of training site information. Section 3.1.3 of the current document provides more details.

Some specific responses:

(a) While we have observed overall accuracies of 85-90 percent in most cases using neural net or decision tree classifiers, the query seems to imply that other approaches could give better results. Except for some isolated and specific studies, the literature does not indicate that other approaches are capable of doing better.

(b) While we will rely primarily on high resolution imagery to identify and label MODIS training pixels, we will also use ground data including plot measurements, wherever possible. Moreover, some types of sites will have multidate imagery, good ground maps, etc.

(c) The advanced classifiers are less restrictive in this area than the conventional maximum likelihood classifier. Heterogeneous classes are recognized as having multiple nodes (neural nets) or leaves (decision trees) and are thus accommodated in the classification process.

(d) Signature extension depends on the quality and quantity of the training data -- ensuring that the training information is truly representative of the population to be classified. We hope that our resources for acquiring training data will be up to the challenge. The EOS test site initiative should help significantly.

10. The panel notes that intermediate variables have been used with simple classifiers to produce good results, and further remarks that single date images have been used effectively.

Intermediate variables have been discussed in point 6. above. They will not be overlooked, nor will simpler classifiers.

Single date images are more problematic, in that a single date strategy would involve classifying many individual images and mosiacing the results. We are not unaware of the success Malingreau's group at Ispra has had in classifying single-date AVHRR images; however, their work focused only on classes of primary, secondary, and non-forest. Note that the heritage of coarse resolution land classification using AVHRR NDVI time trajectories documents the information value present in the multitemporal signal. We don't believe it is desirable to discard this information source in moving to a single-date strategy.

II. Value of data product to the Land science community

1. The panel remarks that the value of the product would be significantly enhanced if produced at higher spatial resolution, and that if at 1-km, all that can be hoped for is better accuracy than pre-launch products.

We recall that it was only three short years ago that we were staunchly defending the product at 1-km resolution in the face of assertions that 1-km was too fine! Point I.3 above responds specifically to this concern. We believe we will have significantly better accuracy than the at-launch product, for two reasons. First, we will have real 1-km resolution, not just a 1-km product produced from multidate data registered to 3-5 km as is the case for AVHRR. Second, we will draw on information from spectral, spatial, and directional domains that are not available to any at-launch products now being developed.

III. Soundness of validation strategy

1. The panel notes that missing data may be a problem.

We agree that this problem needs more attention, although many of the component data streams will be quite robust in this regard (e.g., BRDF). As the panel notes, interpolation (e.g., FASIR) may be a possible option.

2. The panel observes that a water mask will be required.

The EOS/MODIS land water mask is the subject of much attention at present. A 1-km database is currently being assembled combining efforts of both JPL and EDC. It will be built into the cloud mask.

We have also considered the possibility of providing a more accurate post-launch land/water mask as a by-product of the land cover product. It is certainly possible, but to do it depends on the level of effort it will require and the resources available.

3. Reference data will not be perfectly accurate, and the implications of this need to be worked out, according to the panel.

We agree that this subject needs some attention. The statistical literature has some material on this, but we have not applied it specifically to our situation as yet.

4. The panel remarks that our concern about BRDF effects in mountainous regions is well-placed, but we do not provide information about proposed solutions.

Again, this is a subject that needs further attention. The problem may not be overly severe, however, if training sites include mountainous areas and advanced classifiers capable of accepting heterogeneous training sites are used.

5. The panel notes that fuzzy labeling has merit, but that within-pixel proportions will still be required.

Although fuzzy labeling is possible, we don't plan at this time to implement a fuzzy labeling scheme as part of the routine product. However, the mixed pixel issue is important in the accuracy assessment process, and we may try to accommodate it there. Note that the quarter-degree climate modeler's grid product will provide the proportions of land cover types within each grid cell. At 1-km resolution, Aaron Moody of the University of North Carolina, Chapel Hill, has an EOS Young Investigator grant focused on applying neural networks to the proportion estimation problem. He worked on MODIS land cover while earning his Ph.D. at Boston University, and so has close ties with our effort.

IV. Near-term recommendations

1. The panel recommends that one MODIS-based land cover product be generated as soon as possible after launch.

The presumption here is that there is sufficient information in nontemporal domains to provide an accurate land cover classification. This may well be true for many parts of the world, and we agree that it is worth a try. It will also be a good "pathfinder" activity for our processing of the full multitemporal dataset. Another possibility would be to use MODIS information to update the at-launch product in some way, since it will most likely be based on AVHRR data from 1992-1993.

2. The panel recommends that the product have categories that can be categorized and/or generalized for use by carbon, energy, and water cycle modelers and be validated with 100 percent accurate data where possible. The prelaunch product should be validated.

The IGBP categories have been vetted by the IGBP core projects, which include these modelers. We will endeavor to comprehensively validate our product paying due attention to the accuracy of the validation data. The at-launch product will be validated (see response to I.2 above).

3. The panel recommends that we collaborate with IDS teams and others in comparing classification strategies and evaluating the product.

We are currently beginning another round of outreach to IDS teams, and will continue our efforts in that direction. The EOS validation team, to be proposed and selected in the near future, will also provide a source of new investigators interested in land cover data.

MODIS Land-Cover Change

I. Technical soundness of the algorithms/approach described.

1. The panel observes that we are considering multiple algorithms for land-cover change and notes that for the change product, the consistency, stability and "meaning" (information content) of the product should be all relatively stable.

We agree with the latter statement, and are focusing on the change-vector approach as the primary tool for change detection and characterization. However, other algorithms may be needed to recognize and quantify different types of changes, and thus we are also developing these.

II. Soundness of the validation strategy.

1. The panel observes that validation will be difficult, as there is no clear surface variable to compare with the output of the product. They note that the land cover test sites are not selected to validate change, but for training and test of land cover classification.

As the algorithm becomes more refined, it will be easier to develop a specific validation plan for the exact output produced. Our current plan for test sites (see section 3.1.3) includes change sites specifically.

This EPS image does not contain a screen preview.
It will print correctly to a PostScript printer.
File Name : figure.lc.sc.ps
Title : /fs/quack/home/borak/research/modis/LC/ATBD_MOD12/1996/BORAK_REVISIONS/round_3
CreationDate : 9:01 AM December 5, 1996
Pages : 1

Figure 1. Land Cover Parameter process flow

This EPS image does not contain a screen preview.
It will print correctly to a PostScript printer.
File Name : figure.lcc.sc.ps
Title : /fs/quack/home/borak/research/modis/LC/ATBD_MOD12/1996/BORAK_REVISIONS/round_3
CreationDate : 9:01 AM December 5, 1996
Pages : 1

Figure 2. Land-Cover Change Parameter process flow.

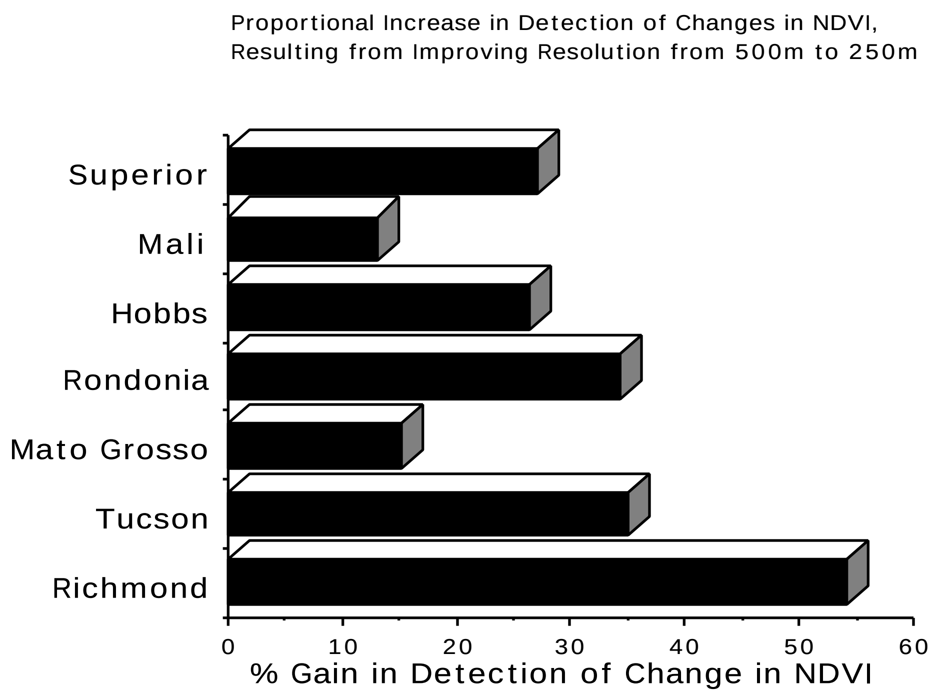

Figure 3. Advantage of a 250m resolution compared with 500m resolution in the detection of change in the normalized difference vegetation index between different dates for seven test areas (Townshend *et al.* 1991).

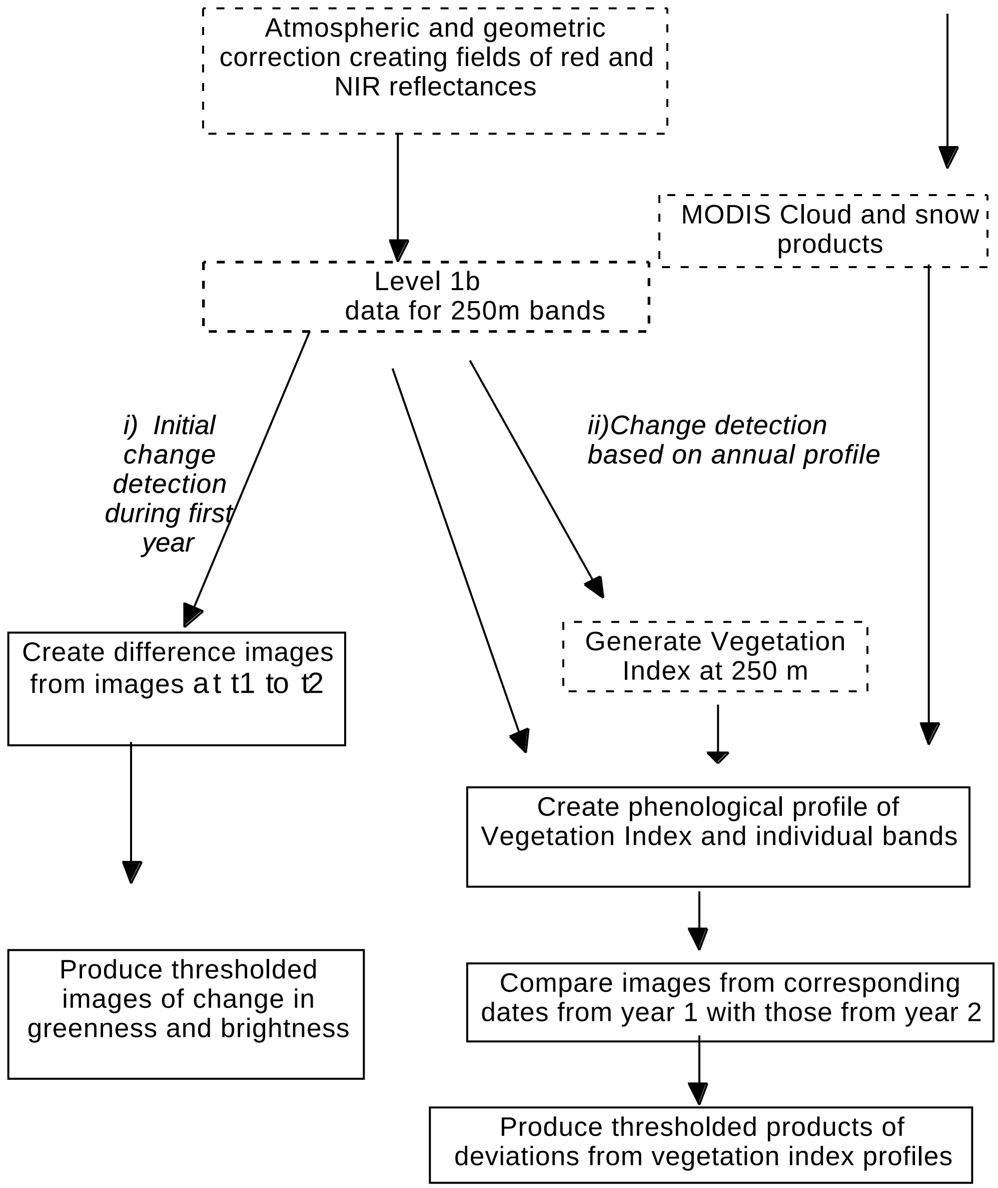

Figure 4. Schematic of the processing chain for the land cover change indicator products. Dashed boxes indicate existing MODIS products.

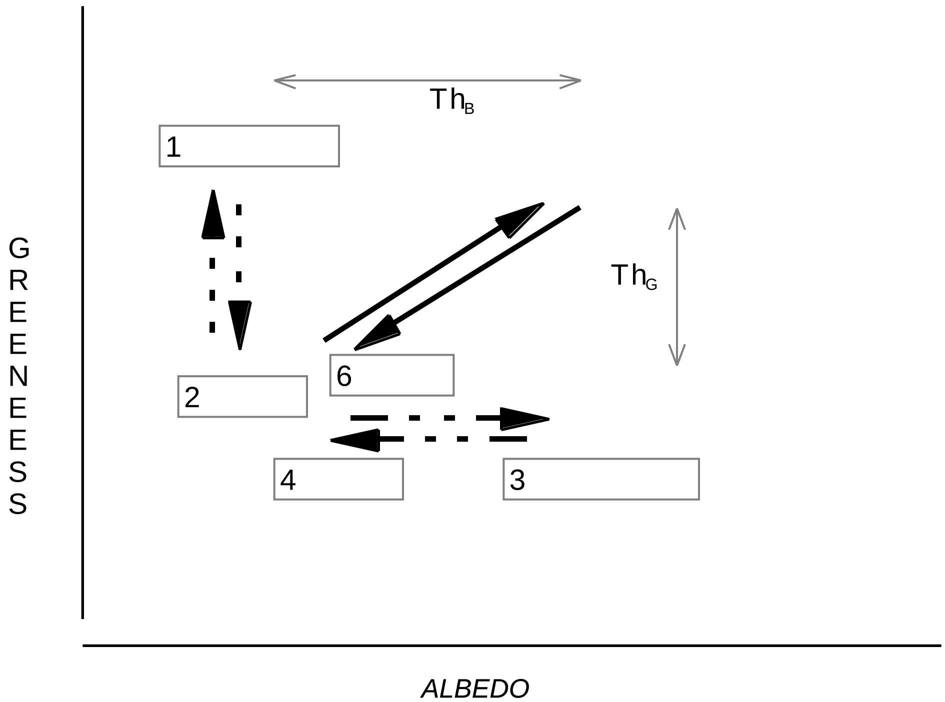

Fig. 5 Thresholds in greenness albedo space.

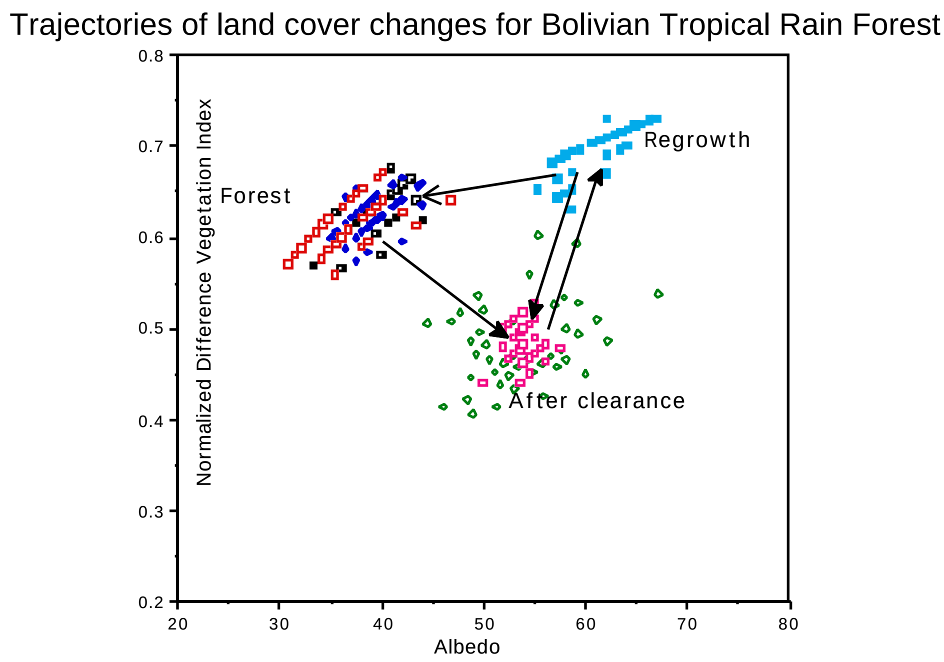

Figure 6. Changes in land cover and the resultant changes in spectral response and the corresponding trajectories in the space defined by the normalized difference vegetation index (greenness) and albedo based on multitemporal Thematic Mapper data for Bolivian tropical rain forest. Following deforestation and conversion to pasture the albedo increases and greenness drops substantially; regrowth is associated with an increase in greenness and a small increase in albedo. Clearance of the regrowth leads to a reversal of the former trajectory. If regrowth is allowed to continue then ultimately the spectral responses of the forest will occur. The solid arrows indicate observed transformations and the open arrow inferred changes.